

nature

8 June 1978

Hazard export: a case for international concern

THERE is an important omission from many of the lists of topics offered for discussion at next year's United Nations Conference on Science and Technology for Development (UNCSTD). This is a subject that has been attracting steadily increasing attention and concern in the industrialised nations, namely the health hazards associated with occupational exposure to a variety of substances used in industrial production.

Recent years have seen the list of chemicals of known or suspected toxicity grow at an alarming rate. Less noticed has been emerging evidence of the gradual transfer of many potentially hazardous production processes from the industrialised nations to Third World countries. Typical examples have been the growth of the benzidine dye industry in India, the shift in asbestos textile production from the United States to Latin and South America, and plans to establish vinyl chloride and arsenic production facilities in South Korea and the Philippines respectively.

Such shifts in location have taken place for a number of reasons. These include the existence of cheap and abundant supplies of labour power, and frequently attractive tax incentives. An additional factor—often quoted by multinational companies—has been the increasingly stringent controls placed on these activities by industrialised countries.

The net result is that potentially hazardous production processes are being shifted to countries which are often less aggressive in their efforts to protect the health of workers, or where the pressures of unemployment (and the rewards of a viable export enterprise) are such that well-intentioned efforts in this direction are constantly thwarted. According to a recent study, for example, certain areas of India have experienced a four-fold increase in the incidence of bladder cancer directly related to the production and use of synthetic dyes; some factories were still found to be using betanaphthylamine, one of the most potent carcinogens known to man, which has been banned in several Western countries (and initially in the UK) for the past ten years.

It is sometimes claimed that concern over occupational hazards in developing countries is misplaced; that the need for income is often so great as to make a certain level of exposure to hazardous substances more acceptable than it would be elsewhere. Yet this is as short-sighted as it is morally suspect. Short-term financial gains must be weighed against the social costs of medical services and welfare benefits required as a result of occupational health damage, as well as against the unquantifiable human suffering caused; neither are conventionally included in the 'costs' of production.

The exact size of the current problem is difficult to

estimate, largely because medical and health records in many developing countries provide inadequate data for reliable epidemiological studies to be carried out. But there is already sufficient evidence to indicate that the situation deserves the attention not merely of international medical bodies—many of whom are already aware of aspects of the problem—but all those concerned with global patterns of technological development.

As with so many similar problems, diagnosis is simpler than cure. But a number of suggestions can be made. Clearly a top priority is that each nation should develop and enforce its own safety regulations, taking what steps it can to protect the health of its workers. Beyond this, however, there are various actions which might be taken at the international level.

The most effective step—although also the most difficult to accept as a realistic proposal—would be for international agreement to control the development of industrial processes known to present potential health hazards, possibly through economic and trade sanctions. However it is one thing for a country to adopt regulations that protect the health of its own workers; it is quite another to take measures, possibly at economic cost, to protect those in other countries.

More realistic in the short term would be international cooperation to ensure that full information is available to all concerned about potential hazards. Barry Castleman, the author of a report *The Export of Hazardous Factories to Developing Nations*, which has recently appeared in the US and is soon to be published in the *Congressional Record*, has suggested an international "hazard export information service" to fulfil this function.

The industrialised nations might make the publication of such data a requirement for granting export and import licences (a suggestion already being pursued, although without much enthusiasm, by the US administration on domestically-banned products such as Tris-treated nightwear). And in the developing countries, careful records of imported technologies and materials would, if nothing else, help to identify population "at risk" as targets for careful observation.

Mere information, of course, is insufficient to guarantee the appropriate action. This requires political will. And here some international declaration—possibly agreed at the UNCSTD meeting comparable to that accepted last year on toxic chemicals by the United Nations Environmental Programme—while not necessarily in itself achieving very much, might still provide a base from which those concerned with protecting the health of workers in developing nations could articulate and defend their demands. □



Why we should ban the military use of satellites

Artificial earth satellites are becoming an integral part of the nuclear arms race. Bhupendra Jasani, of the Stockholm International Peace Research Institute, argues for a new Outer Space Treaty to control the military use of satellites.

THE PEACEFUL uses of satellites cannot be disputed. They have proved to be of enormous value in communications, meteorology, cartography, navigation and in verifying arms control agreements. The military use of artificial satellites, on the other hand, has not only been firmly established but has become a part of the arms race. The tendency now is to turn outer space into a battlefield of the future.

In the 20 years since the start of the space age, 1,482 military-oriented satellites have been launched, which represents about 75% of all satellites. Of these 1,482, about 40% were photographic reconnaissance satellites. The remaining 60% have been used in all kinds of refinements in the technology of fighting wars on earth.

The functions of military satellites range from reconnaissance, navigation or communications to meteorology and geology. They are well on the way to being able to navigate lethal weapons to their targets with a high degree of accuracy; they can predict weather conditions to facilitate bombing; they can be used to pinpoint geographical positions with great precision so that no target can remain in obscurity; they increasingly mean that areas of war can be controlled remotely.

Surveillance and early-warning satellites have become an important part of strategic doctrine. New technology is now available to enhance their performance and to improve their survival chances against enemy attack. With the development of these types of satellites, there seems to be a trend towards a new doctrine of flexible response and limited nuclear war. In this new strategy, space technology can provide better centralised command and control over military forces. In a limited exchange of nuclear weapons it is necessary to assess damage quickly, to make a prompt response and to have precise target information. Reconnaissance satellites provide this capability.

The satellites mentioned so far can be used for peaceful as well as military purposes. But there also exists a category of satellites whose function is purely hostile—the interceptor-destroyer satellites (or as they are often called, hunter-killer satellites). The development by the United States and the Soviet Union of satellites and systems for destroying satellites in orbit, has opened up space as yet another area in which man can wage war. Satellites which can hunt down another satellite and destroy it by exploding nearby have already been tested.

Discussions between the United States and the Soviet Union on banning the hunter-killer satellite begin this month. However, to ban only some military satellites will inevitably lead to the usual criticism of all partial disarmament measures. For example, the 1967 Outer Space Treaty only prohibits placing "in orbit around the Earth any objects carrying nuclear weapons or any other kind of weapons of mass destruction". In Article 4 of the treaty "the establishment of military bases, installations and fortification, the testing of any weapons and the conduct of military manoeuvres" are prohibited only on *celestial bodies*. The treaty has not helped to check the proliferation of military satellites. Therefore, it would seem appropriate to suggest that all military uses of artificial earth satellites be banned.

The fact that the premature re-entry into the earth's

atmosphere of the Soviet Cosmos 954 ocean-surveillance satellite carrying a nuclear powered generator prompted President Carter to say "we would be glad to forgo the deployment of any such satellites altogether and will pursue that option along with the Soviet Union" may be a step in this direction. Between them, the Soviet Union and the United States have launched 38 satellites carrying nuclear power sources. At least four accidents are known to have occurred involving these satellites, two of which were military. In both cases there was radioactive contamination of the atmosphere and the earth's surface. Since military satellites are likely to use nuclear power sources rather than the conventional solar cells in order to make them less vulnerable to attack, it would be desirable, on environmental grounds alone, to ban such military satellites.

Any ban on satellites must be restricted to their military uses. Reconnaissance satellites for verifying arms control agreements and early-warning satellites help to stabilise the relationship between the two main nuclear powers. A step forward in arms control would be to enable more countries to take part in verification which, until 1975, when China emerged as a reconnaissance satellites power, was the monopoly of the USA and the USSR. An international control agency, preferably within the United Nations, should be created which would be the only organisation allowed to launch reconnaissance satellites for checking compliance with disarmament agreements.

This concept of internationalisation could and should be extended to all other satellites. International bodies are already in existence which could control both communications and meteorological satellites. Transmission frequencies used by communications satellites and their orbital data could be monitored by, for example, the International Telecommunications Union (ITU). The World Meteorological Organisation (WMO) could take on the task of controlling the use of meteorological satellites. Controlling navigation satellites could be done by two other UN organisations, the Inter-Governmental Maritime Consultative Organisation (IMCO) and the International Civil Aviation Organisation (ICAO).

There is no appropriate body within the UN to take on the task of controlling geodetic satellites. Outside the UN, however, one of the committees of the International Council of Scientific Unions (ICSU), the Committee on Space Research (COSPAR), could foster international co-operation and control in all scientific experiments that make use of rockets and satellites as research tools. This could include the use of geodetic satellites.

In any Comprehensive Outer Space Treaty, verification would be a difficult but not impossible problem to solve. It is possible from measurements of the orbital characteristics of satellites, the types of signals they transmit and estimations of their shape, weight and dimensions to determine their missions.

A number of nuclear arms control agreements have been reached so far but these have resulted in controlled armament rather than disarmament. At present the total nuclear arsenal of the nuclear-weapon powers is not increasing significantly. However, the quality and effectiveness with which such weapons can be delivered to their targets is improving dramatically. The artificial earth satellites play a major role in this trend. If, by banning the use of military satellites, the qualitative nuclear arms race could be checked, then it would mean that the nuclear-weapon powers would be making a positive step towards fulfilling their pledge made in Article VI of the 1968 Treaty on the Non-Proliferation of Nuclear Weapons. □

University fights secrecy ban on computer research

A COMPUTER scientist at the University of Wisconsin's Milwaukee campus has been warned by the federal government not to discuss the results of his research with colleagues because to do so might infringe national security regulations.

Dr George I. Davida, associate professor of electrical engineering and computer sciences, has been conducting research for a number of years supported by the National Science Foundation into the security of computer-based data systems. Although the research is unclassified, Dr Davida has now been told by the Department of Commerce, with which he had filed a patent application resulting from his research, that the study "has been found to contain subject matter, the unauthorised disclosure of which might be detrimental to the national security".

The department has added that unauthorised disclosure of the study, which would include discussion of its results with academic colleagues, could be a violation of patent laws. These carry penalties of up to two years in prison and \$10,000 in fines.

Officials at the university are stunned by the government's action, which they consider to be a serious intrusion of academic freedom, and are seeking legal advice on how to challenge it.

Research into the security of computer-based data systems has become a controversial area in recent years, partly because of concern among officials of the National Security Agency that the open discussion of research results could prejudice their successful application by the federal government.

Last year the agency was accused of harassing scientists working in cryptography during efforts by the National Bureau of Standards to certify a single data encryption standard to be used for all government non-classified information.

An investigation into the NSA's alleged activities was carried out by the Senate Select Committee on Intelligence, which concluded that "there has been no direct or indirect harassment of scientists working in the field of computer security," (a conclusion which will now no doubt be revised).

However, the committee's report added that some NSA officials had expressed concern to the National Science Foundation about certain grants with cryptological ramifications, and suggested that "NSF and NSA should initiate efforts to reduce the ambiguity and uncertainty which surrounds the granting of research funds for public cryptography".

Professor Davida has received two grants from the foundation. The first, which was awarded in 1975, was for a one-year study of cryptographic algorithms for computer data bases. This work involved developing a theory of computation cryptography leading to algorithms for data security having certain required properties (such as resistance to cryptanalysis).

A second grant of \$89,700 was awarded to Dr Davida last year for a two-year study into the theory and practice of data security. In particular he has been working on models for analysing data-base security in situations where information-system users can sometimes retrieve data they are

not authorised to obtain.

Dr Davida submitted a patent application arising from his work in October 1977. As a result of this application, and comments on it received by the Department of Commerce from "a defence agency", he has now been notified from the Patent and Trademark Office that he must not publish the results of his work or discuss it with colleagues. Furthermore any discussions about the research with "foreign nationals" must be officially reported.

Dr Davida would not comment on the affair last week, but issued a statement saying that "because of unanswered legal questions I am unable to discuss a matter of vital concern to me and my colleagues in the academic community." Other computer scientists are surprised the government has reacted so strongly.

Dr Warner A. Baum, chancellor of the University of Wisconsin (Milwaukee) said he was "stunned by the notion that a university professor should get an order from the US government saying that he cannot talk to people about his research". The university does not allow classified research on any of its campuses following bombing of a mathematics building on its Madison campus during anti-Vietnam war demonstrations in 1970.

A spokesman for the NSF said that the foundation had not yet decided whether to join the university in challenging the secrecy order. The case, and its implications for academic freedom, were being given urgent consideration.

David Dickson

Lead levels in Birmingham are typical of urban areas, says UK report

THE results of recent lead pollution studies in Birmingham have focused attention in the UK on the long-standing controversy over the neurological effects of low lead levels in the body. Although the study found that blood lead levels in its sample of Birmingham people were not unusual for urban dwellers, the almost simultaneous release of the results of studies in Germany and the US, suggesting that even these levels could be responsible for mild neuropsychological disturbance in children, have caused some concern.

This is not the first time Birmingham has been the focus for this debate: four years ago a study of blood and

Barnaby

Spaghetti junction: a clean bill of health

air lead concentrations in the area around a new motorway interchange in part of the city suburbs aroused some public concern. The latest study, carried out by a joint working party set up by the UK Department of the Environment in the wake of that concern, set out to assess how the presence of the same motorway system at Gravelly Hill, known colloquially as Spaghetti Junction, has affected local lead levels after several years of

Sorry, for copyright reasons some images on this page may not be available online

operation.

The working party found that after the interchange opened, air lead levels near it increased—though not as much as the increase in traffic. It attributes the unexpectedly low increase to the elevated and well-ventilated position of the interchange. However, levels near the junction were found to be higher than those measured at other sites in the city but still below the maximum level currently proposed by the Com-

mission of the European Communities of an annual average of monthly mean values of $2 \mu\text{g}/\text{m}^3$.

Measurements of blood lead levels of adults and schoolchildren living in different parts of the city showed no unusually high values for urban dwellers and no significant increase in those living near the interchange. In the case of the children, however, those living in the inner city areas had slightly higher values than those living in outlying districts. No child however had a level higher than $35 \mu\text{g}/100 \text{ ml}$, the maximum recommended by the European Commission.

The results of the study of pre-school children (aged one to four) are more alarming. Seven out of 83 examined from the inner city areas had levels greater than $35 \mu\text{g}/100 \text{ ml}$. None of them showed any significant psychological disturbance, however, and in only one case could the source of lead be identified—lead rich dust in the home from the preparation of old painted surfaces.

Dr H. A. Waldron, one of the working party members currently at the London School of Hygiene and Tropical Medicine, is extending the work on pre-school children by carrying out another study using a different sampling method. So far his results are in line with those of the original study.

Sorry, for copyright reasons some images on this page may not be available online

Calm beneath spaghetti junction

The report of the working party's study effectively gives the Gravelly Hill interchange a clean bill of health. However, Dr Robert Stephens, of the Department of Chemistry, University of Birmingham and also a member of the working party, feels that even the 'normal' levels of blood lead found could be affecting the central nervous systems of as many as 20% of school children.

He has based this figure on the results of recent studies in Germany, Canada and the US which have attempted to correlate minor, clinically non-specific psychological disturbances with lead levels in blood and teeth. The US and German studies found

disturbances in children with lead dentine levels of $20 \mu\text{g}/\text{g}$ and $13 \mu\text{g}/\text{g}$ respectively. A previous study of whole tooth lead in Birmingham children showed that 20% had levels higher than $15 \mu\text{g}/\text{g}$. Hence Dr Stephens figure.

Not all of the members of the working party are as concerned as Dr Stephens about the possible implications of these studies. Different methods of measuring blood and dentine lead levels can give widely varying results, according to Dr R. Archer, Chief Officer in Birmingham's Environmental Department. Even the working party had difficulty at first, he says, in getting their measurements to agree with those taken at the same time by the Institute of Child Health in London.

The present working party is being disbanded soon but Dr Archer will be setting up a new steering group within the next few weeks to decide what subsequent studies should be done. Psychiatric tests in conjunction with measurement of levels of lead in the body are not the only possibilities, however: other areas which the current study has thrown up as needing further investigation are the significance of various forms of lead intake, from food or water for example, and the ways in which lead can be transported from say the work place to home.

Judy Redfearn

NSF budget under fire

FOR the third year in succession a key US congressional committee has recommended major reductions in the research budget requested by the National Science Foundation. The House Appropriations Committee has accepted a subcommittee recommendation that \$893.9 million be appropriated for the NSF in the fiscal year 1979, in contrast to the \$934 million requested by President Carter.

The president's request has already been accepted virtually in full by both House and Senate authorisation committees; in contrast the House Appropriations Committee's recommendation would mean that the increase in the foundation's budget would be reduced from 10% to 3.8%, considerably below the cost of living. In particular, the committee suggests that the NSF's research programmes and related activities should be given a budget of \$806.4 million, which is \$44 million behind the administration's request, allowing an increase of 3.0%.

The main target of the committee's criticisms is the foundation's newly established Applied Science and Research Applications Division (ASRA), which was set up at the beginning of the year as a successor to the widely-criticised Research Applied

to Social Needs (RANN) programme. Stating concern, for example, that ASRA's integrated basic research programme covers "precisely the sort of relevant basic research for which the mission-oriented agencies are responsible," the committee suggests that the programme be limited to \$48.1 million in 1979, rather than the \$73.9 million requested.

At present, NSF officials are not too concerned at the committee's actions. Previous attempts by the same committee have been resisted by the Senate, and have been at least partially restored at the conference stage; it is hoped that the same will happen this year, particularly in light of the president's stated awareness of the problems of university research.

One positive aspect of the appropriations committee's proposals is to increase the administration's request for science education from \$77.6 million to \$82.6 million.

David Dickson

Argentina pressed to free jailed scientists

URUGUAY appears to have taken "a first step" towards restoring human rights, but Argentina continues practices which are "clear violations" of

human rights, according to a delegation of the US National Academy of Sciences, which recently visited them.

The delegation (NAS members Christian B. Anfinsen and Robert P. Perry) reported that in Uruguay they were able to visit the mathematician José Luis Massera who has been in prison since 1975 for his political views. The NAS is at present petitioning the Uruguayan government to allow Professor Massera and his wife to be exiled to a country willing to receive them. The jurist in charge of the case has indicated that "the plea and presence of the NAS could be viewed in a positive light".

In Argentina, however, the delegation found it impossible to visit a political prisoner who had been offered a research post at a US university, neither was the Minister of the Interior willing to meet them to discuss their concern about political prisoners and *desaparecidos* ('vanished persons'). As a result of the delegates' report the NAS Committee on Human Rights is petitioning the Argentinian government to release Elena Sevilla, a physicist, who is still in custody on political charges of which she was acquitted in January 1976, and Dr Claudio Santiago Bermann, a psychiatrist, who is being held in preventive detention without formal charge or trial. □

Vietnam extends scientific links

IN March/April of this year, a four-man scientific delegation from the World Federation of Scientific Workers (WFSW) visited Vietnam. Professors E. H. S. Burhop (University College, London), J. M. Legay (Secretary-General, WFSW, France), A. Matveef (Moscow University), and E. W. Pfeiffer (University of Montana) spent most of their ten-day stay visiting scientists in research institutes in both Hanoi and Ho Chi Minh City. A draft statement issued by the delegation reports that the Vietnamese scientists require as a matter of some urgency scientific literature; funds to purchase research equipment and supplies; and more contact with scientists abroad.

The delegation was impressed with the importance attached by the Vietnamese to science and technology. The object of the visit was to assess the scientific development of the country. But the group failed to elicit information on damage caused by the defoliation programme. Professor Pfeiffer, the US member of the delegation and a zoologist, is widely known for his campaign against the use of herbicides during the war in Vietnam. Pfeiffer was one of a handful of US scientists who campaigned relentlessly on this question and were instrumental in making it an issue of public debate. He told *Nature* that it was with some disappointment that he found the Vietnamese unwilling to take him to areas affected by herbicides; he points out their unwillingness to reveal information relating to the effectiveness of the spraying programme for fear that it could be used by the US in perfecting

its chemical weapons.

As regards US scientists helping their Vietnamese colleagues, Pfeiffer feels that there is little that can be done by the individual "other than to work for the lifting of the embargo imposed by US Congress against trade with Vietnam and other essentially political activities". He mentions the delegation's visit to the IBM Computer Company in Ho Chi Minh City, where there are several million dollars worth of computers in operation at the moment. Spare parts and repairs will, however, be a problem in a year or so, and the Vietnamese are keen to establish relations with IBM. If this is not achieved, then Vietnam will have to turn elsewhere for assistance, probably to the socialist world. Pfeiffer therefore considers it to be most important that this fact be drawn to the attention of the politicians in Washington. He says that if the embargo continues for much

Sorry, for copyright reasons some images on this page may not be available online

US plane spraying defoliant, 1966

searchers were leading. They assigned the subjects or the methods (they often provided the devices themselves)." Gradually, said Khristov, the Bulgarian contribution has become greater.

Speaking at the National Conference of the Bulgarian Communist Party, Academician Angel Balevski, President of the Bulgarian Academy of Sciences recently observed that co-operation within Comecon, first and foremost with the Soviet Union, will continue to be the mainstay of Bulgarian science policy.

At present, he said, priority is being given to research aimed at mastering scientific and technical progress and its transfer to production—a trend reflected, incidentally, in Bulgaria's plan and budget for 1978 which allots 909 million leva to introducing scientific and technological developments into the economy, including 790 million leva for capital investments, and 450 million leva for the logistics of implementation.

longer, the result would be that the Vietnamese would be forced into a political stance which America "spent billions [of dollars] for armaments to prevent".

● India is one of the most recent countries to enter into a scientific aid agreement with Vietnam. The agreement, signed on 26 February this year during the visit to India of Pham Van Dong, prime minister of Vietnam, provides amongst other things for the exchange and training of scientists, the establishing of joint research programmes and the exchange of scientific and technical information. Some of the programmes will probably concern rice research, buffalo breeding and water resources management.

● Britain, too, appears to be improving its relations with Vietnam. On 2–15 May, Professor Ton That Tung, director of the Viet Duc Hospital in Hanoi, visited the UK as the guest of the Foreign and Commonwealth Office. Tung—one of Vietnam's leading medical practitioners, a surgeon researching into the treatment of liver carcinoma—met numerous scientists in the UK for discussions about his research interests. He told *Nature* that he hoped that his visit would lead to greater contact between scientists in Vietnam and the UK.

● THE report of a five-man UN mission to Vietnam has been published. The team visited the country after the General Assembly's call for help in reconstruction. While providing facts and figures on the state of the country, the report offers little new information on damage to forests, and its section on the need for scientific and technological assistance is all too brief.

Alastair Hay

Bulgaria to shake up its scientific cooperation

INTEGRATION and cooperation within Comecon—a basic aim of long-term Comecon planning—is undoubtedly a more tempting prospect to a less-developed country such as Bulgaria than, say, Poland, where in 1976 the leading foreign affairs journal *Sprawy Miedzynarodowe* published some sharp criticism of this policy. But such co-operation holds a certain danger that the less-developed partner will become a mere passive recipient.

The Bulgarians, however, seem well aware of this hazard. Recently, reviewing joint Soviet-Bulgarian work in nuclear physics, Academician Kh.Y. Khristov, director of the Nuclear Research and Nuclear Power Institute of the Bulgarian Academy of Sciences observed "in the interests of truth" that "in fact, our participation, at least in the beginning, was not adequate. Usually the Soviet re-

This trend, Balevski implied, will continue; as also will cooperation, since "a small country like Bulgaria cannot undertake the development of all scientific spheres."

He hinted however, that the development of research will in the future be more "selective"; while cooperation will be aimed at the "quality" rather than quantity of projects.

Academician Balevski did not specify what these projects will be. However, at the opening last week of an exhibition in Moscow commemorating the Bulgarian war of independence, Mr Kosygin, former Soviet premier, stressed that existing plans call for close cooperation between the Soviet Union and Bulgaria in energy, ferrous metallurgy, machine-building, chemistry, electronics, agriculture, light industry and the food industry. Scientific cooperation, it may be assumed, will be closely linked to these broad economic aims.

Vera Rich

Patent rights: once more around the block

WHO is entitled to the rewards of a scientific experiment that results in a technological moneyspinner? The research worker involved? His or her supporting institution? The entrepreneur who provides capital to bring the product to the market place? Or the government agency who supported the research? The debate is not new—indeed it has been going on for at least 30 years, ever since the rapid post-war expansion of federally sponsored research in universities which is now estimated to be worth over \$3.5 billion a year.

The latest round of skirmishing has been sparked off by the recent publication of a planned change in federal regulations. It proposes that universities should, under certain defined circumstances, be automatically granted the patent rights to the results of research carried out on government funds.

At present, different government agencies have different policies on patent rights. The Department of Energy, for example, has inherited from the Atomic Energy Commission a statutory 'title' to the rights on research that it has paid for, although it can decide to waive this right and give it to the university which has carried out the research.

In contrast, the Department of Health, Education and Welfare has no statutory responsibilities, but as a matter of policy is prepared to waive rights if a petition to do so is received. Partly as a result of administrative problems in pursuing this policy (last summer, as part of a review of patent policy, a block was put on further waivers holding up about 30 patents which universities had applied for) the department also has a scheme of institutional patent agreements (IPAS). Any institution which receives new research funds can apply for such an agreement. To qualify, the institution has to demonstrate that it operates an effective technology transfer programme. The IPA, if granted, provides that the institution will automatically receive the patent rights on the research it carries out.

At present the only other agency to use IPAS is the National Science Foundation which has agreed them with 19 universities since 1974. Under the new proposals, a revised form of IPA would be extended to all government agencies sponsoring university research, and would become the standard mechanism for dealing with patent rights.

Both the universities and the agencies are, in general, in favour of

David Dickson reports on the latest skirmishing in the US over who should get rich out of money-spinning inventions

this scheme. Despite reservations on details, both see it not only as a way of cutting down administrative paperwork and promoting standardised procedures, but also as encouraging the transfer of technology to private industry.

What has concerned others, however, and has in particular attracted the attention of the Monopoly and Anti-competitive Activities Subcommittee of Senator Gaylord Nelson's Select Committee on Small Business, is whether, by allowing universities such

cational associations recently, for example, Dr Thomas Jones, vice-president for research at Massachusetts Institute of Technology, said that the total university income from patent royalties was only about \$9 million a year, and that few universities operated licencing programmes in the black.

It was also made clear by Dr Jones, however, that the gain to universities of negotiating licences directly with private industry is not primarily financial, but that such practices help to establish links between the two sectors. Industry is more prepared to collaborate with a university if it can be sure that the results of research will be available for licencing.

The universities will have a further chance to argue their case at a second series of hearings planned by the sub-



"I want to patent my scheme to unravel government patent policy!"

rights, the government is giving away more than is in "the public interest" considering the size of its expenditure on university research. Ralph Nader's health research group, for example, in a move which led to the temporary postponement of the implementation of the new regulations, complained to the General Services Agency that the regulations would permit universities and commercial enterprises to "reap hundreds of millions of dollars of profits from work supported by the federal government". In particular such arguments have recently been made about research involving recombinant DNA techniques, for which Dr Donald Fredrickson, Director of the National Institutes of Health, subsequently decided that no exemption from the standard IPA procedure need be made.

Universities dispute the allegations that patent licencing denies the federal government a major source of income. In particular they claim that the financial benefits received as royalties from the licencing of patents is often much lower than imagined. Addressing hearings of the Senate Subcommittee on behalf of a wide range of edu-

committee for 20, 21 and 26 June. These hearings are expected to focus on ways in which the subcommittee feels that the proposed IPAs would be more liberal than those currently in force.

Two issues in particular have already attracted the attention of the subcommittee, which is at present carrying out a two year study of federal patent policy. The first is the amount of time for which a university will be permitted to grant exclusive licence to a manufacturer who is only prepared to undertake production on this basis. At present, exclusive licences can be granted for a period of eight years, or three years from the date of first sale of a product, whichever is the shorter. Under the new regulations, the three years would be extended to five, and allowance would also be made for time taken by a regulatory agency to make its pre-market clearance.

The universities argue that they must be able to maximise the attraction of a particular patent to a potential investor. "Market development costs are often much greater than the physical cost of invention. Private

enterprise therefore has to be given some incentive in order to be confident of a reasonable profit", according to Mr William Bremer of Wisconsin University.

The Senate Subcommittee, however, has expressed concern at the 'monopoly' effect of issuing exclusive licences. It claims that although under both the present and projected form of the IPA, these are intended to be the exception rather than the rule, in practice most licences are granted on such a basis. A second concern is over who should decide whether a research worker should be granted a licence to develop a discovery he has made. In the initial draft of the new regulations, it was suggested that, in line with current practice at the NSF, this would require the agreement of the agency which had sponsored the research.

When this draft was sent out for comment, however, many universities reacted strongly against it, claiming that it would reduce the incentive for the individual research worker to report a potentially patentable discovery. The provision has now been dropped from the proposed regulations, raising the subcommittee's concern that universities may offer exclusive licences to a research worker rather than seeking outside, more appropriate, licensees.

Are there any alternatives to IPAs? William Carey, executive officer of the American Association for the Advancement of Science, together with others, feels that they should be dropped altogether as an unnecessary impediment on the flow of scientific ideas into the technological marketplace.

Subcommittee staff members are also looking at the possible appropriateness of a government agency such as the UK's National Research Development Corporation, which has a statutory responsibility to handle the patenting and licensing aspects of research carried out on government money including that funded in universities by the UK's research councils.

But at present any major innovation in patent policy, apart from the introduction of the new regulations, seems unlikely. Recognising the complexity of the issues—economic, political and legal—and the intensity of the motions that they raise the White House, for example, has steered clear of the area and it will not be part of a multiagency study of innovation announced recently. "The debate over government patent policy is a thicket a prudent man hesitates to enter," Mr Charles H. Herz, general counsel of the NSF, told the Senate Subcommittee. So it has been—and so it promises to remain.

David Dickson

Friedrich Wöhler RIP



THOMAS H. JUKES

THIS year is the sesquicentennial of the announcement by Wöhler that marked the birth of synthetic organic chemistry. Earlier, Berzelius had formulated a concept that the numerous carbon compounds found in natural sources could be produced only by living organisms. "Elements in the organic world," said he, "appear to obey laws different from those in the inorganic world." His concept was the molecular equivalent of the dogma of spontaneous generation of life, a theory that had been put to death by the Abbé Spallanzani in the eighteenth century. As a result, the theory was, as Voltaire said, "rejected by whoever has studied nature at all". In like manner, Wöhler refuted Berzelius' theory by preparing urea from ammonia and carbon dioxide. In a sense, Berzelius' 'organic' chemistry was no longer organic. But the name persisted, and, as everyone except 'organic food' cultists should know, the term 'organic' properly refers to carbon compounds rather than to vegetables grown with the aid of animal manure.

The vitalists were not extinguished by Spallanzani, and, despite Voltaire's dismissal of their cause, they returned in 1858 to proclaim that spontaneous generation of animals and plants could be demonstrated in the laboratory. Pasteur promptly turned back their challenge with a series of famous experiments. One would have thought that, by 1978, Wöhler's discovery would have convinced everyone that 'natural' compounds can be synthesized. But mythology is more persistent than veracity. Even the use of Wöhler's own compound, synthetic urea,

which is used as a fertilizer, is challenged as not being equivalent to animal excreta for this purpose. The challengers have something to sell—a line of 'natural food' products. The selling job has been so pervasive that an extensive opinion survey in 1972 showed that 65% of the general public, and 52% of college graduates said "No" to the question, "Are man-made vitamins just as good as natural vitamins?"

Recently, at a meeting, when I pointed out that a synthetic vitamin is identical with its natural counterpart, a prominent *Washington Post* reporter said, "Your colleagues don't agree with you". To which I responded that, in this case, my colleagues would be wrong, and I asked her if she had ever heard of a mixed melting point. Of course she had not, and she was obviously not interested in my attempt to obfuscate the discussion by trotting out such scientific jargon.

'Natural', in terms of food, means a blessed surcease from hyperactivity in children, according to true believers. The proponent of this belief, Dr Ben Feingold, tells his followers that the statement "that the whole body is chemical; that all components of food are chemical" is "a somewhat ridiculous public relations argument of the additive industry," also that "not a single one of the synthetics used in our food has been subjected to the rigid investigations required for licensing drugs," which is quite untrue, because synthetic vitamins are used in both foods and drugs. The dietary treatment that he prescribes is based on elimination and exclusion of synthetic substances from the list of permitted foods. This procedure is a curious echo of the ancient ritualistic casting-out of demons from those who were possessed (and therefore hyperactive).

And 'natural' means profits. An officer of the Federal Trade Commission wrote recently of manufacturers who "do have a deep concern with their right to describe products as natural under the conditions described". I suspect that this deep concern may be connected with hope of sales. A vice-president of a large food company told me last month of his company's plans to test-market a new breakfast cereal which, he said, would be really natural, and contain "only natural vitamins". My attention wavered, and then vanished. But perhaps I should send him a copy of *Ann. Phys.* 12, 253 (1828) so that he can read about making urea without kidney, bladder, or lamp-post-seeking possessor thereof.

correspondence

The lipid hypothesis

SIR,—John Rivers' article (3 November, page 2) on the lipid hypothesis may have confused readers not familiar with this subject because it conflicts with the strong consensus of recent scientific opinion. An alternative summary and explanation may be of interest.

Since 1968 no less than 19 completely independent national working parties from many different countries have between them reviewed all aspects of relevant research and reached remarkable agreement. Each (19/19) has recommended reduction in the consumption of saturated fat either by the population as a whole (15) or by those detected by screening to be at high risk (4). Since no country has found it practicable to carry out screening on a large scale, and in any case most of those at high risk could not be detected by present means, it follows that the advice is applicable to all who wish to take reasonable precautions to preserve their health. Most (17/19) have specifically advised partial substitution of saturated by polyunsaturated fat and 17/19 restricted consumption of cholesterol.

Always in human affairs, no matter how strong the majority view, a few dissident voices will be heard, and of course their viewpoints, if they have substance, must be considered. This has been done.

It is not correct to state, as Rivers does, that the DHSS (COMA) panel (*Diet and Coronary Heart Disease*, HMSO, 1974) and the Joint Working Party (JWP) of the Royal College of Physicians and British Cardiac Society (*J. Roy. Coll. Physicians* 10, 213, 1976) reached diametrically opposite conclusions. Both made the same recommendation for the whole community to reduce saturated fat. The difference was that, whereas the DHSS did not suggest what should take the place of saturated fat, on the grounds there was no proof that benefit would result, the JWP, in conformity with the consensus, did. There is of course no 'proof' because, as explained, dietary trials are impossible to carry out. However, as Professor Truswell, who is now a member of COMA, has pointed out, one cannot simply advise a reduction of something without indicating what should take its place. The two obvious sources, both of which are suitable, are polyunsaturated fat and carbohydrate (preferably unrefined) in the form of bread, cereals and potatoes. COMA had no objection to polyunsaturated fat. Consequently there are no practical differences between the two reports.

Rivers supports George Mann, a well-known heretic and none the worse for that, but his criticisms of the consensus are easily countered. The facts are formidable. As emphasised by WHO, mankind is likely to experience the greatest epidemic ever faced unless action is taken. In the UK, as in the USA, one man in three or four is likely to suffer from arterial disease before

retirement and often long before. Coronary heart disease (CHD) is the commonest cause of death in men over 35.

The steep rise in incidence in many western countries in recent years, the wide population variations that cannot be explained by race, the experience of migrants from low to high incidence countries and the changes which occurred in war indicate that the causes must largely be environmental and due to behavioural change. The evidence also indicates that, although other risk factors may contribute, radical changes in our food structure and eating habits are the fundamental cause.

The Japanese are a modern, industrial, stressful, heavily-smoking and unusually hypertensive society, but have little atheromatous disease. This can be explained by their low (normal) blood lipid levels, related to their habitual diet. If Japanese migrate to the USA, however, as they have done in large numbers, and change, in particular, their food habits, the blood cholesterol levels rise, arterial disease develops and in due course they have a similar incidence of CHD to their host country. In recent years similar dietary changes have been occurring in Japan, with the expected results.

The links between diet, plasma cholesterol concentrations, atherosclerosis and CHD are immensely strong. They are based on epidemiological correlations, clinical observations, pathological changes, the vast International Atherosclerosis Project which correlates diet, blood lipids and arterial changes, laboratory studies on mechanisms and on experimental work. If man's nearest relatives, the non-human primates, are given a similar human diet, they develop a similar rise in blood cholesterol and, in due course, similar arterial changes with similar complications. Encouragingly, if those who have developed arterial disease return to their natural diet, considerable regression can occur.

The final link by controlled dietary studies in man is impossible to forge owing to the practical difficulties. It is not correct to say that dietary studies are being carried out that could bring final proof. Mann makes much of inconclusive dietary studies, but it is now appreciated that although the matter had to be put to the test (and most think the results were encouraging), absolute proof cannot be obtained. Under these circumstances, action has to be based on the best available evidence.

Rivers refers to dietary studies within populations and refers to Michigan as an example. In affluent communities, however, most people habitually consume a similar rich diet, with relatively high serum cholesterol levels compared with low incidence countries. The very high correlations which have been obtained are from studies in countries with widely different habits, and they are very convincing.

It is unfortunate that Rivers should

again bring up the cancer scare. This was raised some nine years ago in an elderly population being studied in Los Angeles. The matter was immediately investigated but not confirmed either in that study or in any other. As Heady, from the London School of Hygiene and Tropical Medicine, wrote in the *British Medical Journal* (1, 115, 1974), "the case is not merely unconvincing, it does not stand examination".

The unsaturated vegetable oils now being recommended are consumed in similar amounts by many European populations with a very low mortality from CHD and cancer. They are also consumed in large amounts in South Russia. The question of *trans* unsaturated fatty acids has also been very thoroughly studied. No harmful effects have been found experimentally and they have been incorporated in margarines for about 70 years without untoward effects. Their concentration is higher in butter than in polyunsaturated margarine.

It is also incorrect to state that over 90% of body cholesterol is synthesised. The body synthesises sufficient cholesterol for all its needs and all dietary cholesterol is surplus to requirements. However, isotope studies have shown that this amounts on average to about 25% of plasma cholesterol, and in some it is more (Kaplan, J. A. *et al.*, *Arch. Path.* 76, 359, 1963). This is a substantial 'extra'. There is considerable individual variability, but it could be critical. Another important fact is that in response to cholesterol consumption, tissue concentrations may rise steeply with little or no change in the plasma. The permeability of the arterial wall is an important variable factor, but unfortunately cannot be studied by simple laboratory tests.

As regards the Masai, the relevant fact is that they have very low serum cholesterol levels which would account for their freedom from CHD. As Mann himself showed, any atheroma present is mild and, even if the vessels had not become widened, there would be no obstruction.

As regards exercise, the East Finns, who have long had the highest serum cholesterol levels in the world and the highest mortality from CHD, habitually undertake more exercise than the Masai. If the diet is faulty, regular strenuous exercise is clearly not protective. This is not to deny the value of exercise in reducing the risk of CHD. There is some evidence of this even in Finland.

Doubtless new information on mechanisms will be obtained by future research, but it is very unlikely it will influence the need to correct some of the radical changes man has made in his food structure and eating habits in recent years. The principal problems are not what to do or how best it could be done, but those of motivating the many concerned.

Your faithfully,

R. W. D. TURNER

University of Edinburgh, UK

news and views

Somatic mutations

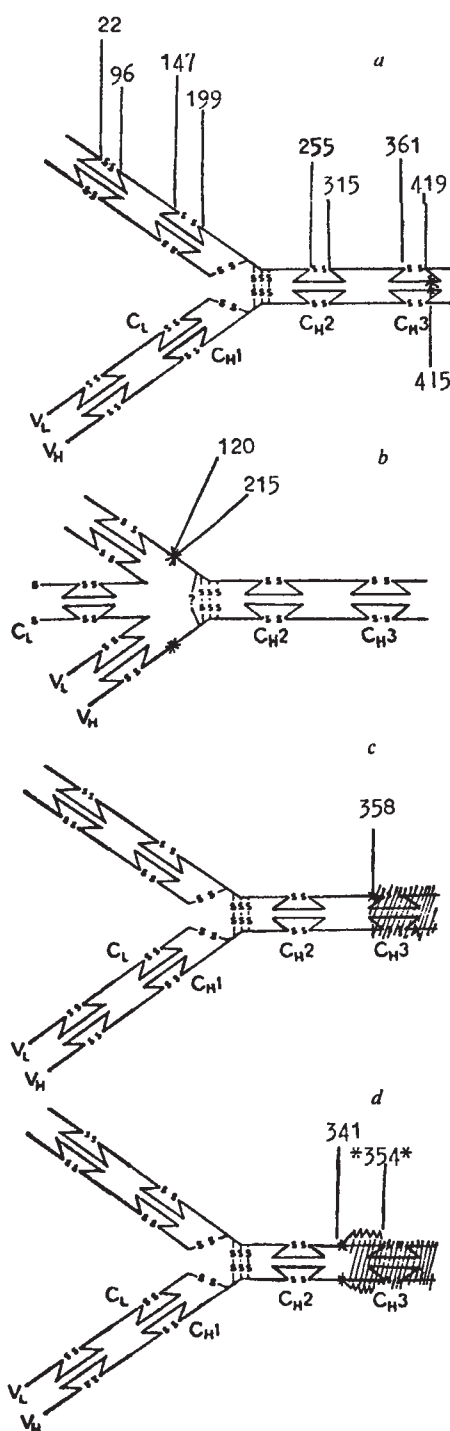
F. Cramer

SOMATIC mutations in eukaryotic cells have been invoked as the source of antibody diversity, the underlying cause of cancerous changes, and of the changes that occur during ageing. However, relatively little is known at the molecular level about the type and origin of mutations in mammalian cells, compared with bacteria.

One of the most intensively studied systems is that of the clonal variants of the mouse myeloma MOPC-21 cell line, in which C. Milstein and his colleagues have characterised a series of spontaneously arising variants of the immunoglobulin gamma heavy chains.

The exact structure and sequence of the wild-type gamma protein synthesised by MOPC-21 are known, as is the sequence of the mRNA coding for this protein. Spontaneously arising variants have been isolated, cloned (see Secher *et al. Immun. Rev.* **36**, 56; 1977) and systematically analysed by fingerprint analyses of the proteins and mRNAs, cell-free translation of the isolated mRNAs, and by cell fusion complementation analyses. To date it has been possible to identify and study four spontaneous structural mutants, designated IF1, IF2, IF4 and most recently IF3 (Adetugbo & Milstein *J. molec. Biol.* **121**, 239–254; 1978); and a number of non-secreting and/or non-producing variants (see Secher *et al. op. cit.*) in some of which an inactive mRNA was detected.

Concentrating on just the four spontaneous structural mutants, several major points can be made. The three mutations resulting in deletions (IF1, IF2 and IF3) occur outside the pseudosubunit disulphide loops of the gamma heavy chain even though more amino acid residues occur inside the loops (247) than outside (193). The structure of MOPC-21 Ig, the pseudosubunits or domains, the positions of the mutations and the resulting structures of the variants are indicated in the figure. Both IF1 and IF3 have



a major portion of the C_{H3} domain deleted. In IF1 a 'nonsense' mutation occurs at amino acid residue 358, resulting in early termination. In IF3 a two-base deletion at residue 341 causes a frameshift resulting in an aberrant sequence continuing to residue 354 at which a premature 'ochre' termination codon occurs. It is interesting that two of the four spontaneous mutants have errors arising within a 20-residue region between the C_{H1} and C_{H3} domains. IF2 is an intracistronic deletion and occurs at the C-V junction deleting C_{H1}. The mRNA for IF2 gamma-chain is smaller than the wild-type and the other variant gamma protein mRNAs. Many of the recorded human heavy-chain disease syndromes show identical deletions in the gamma chains of patients (Franklin *New Engl. J. Med.* **294**, 531; 1976). The change in the IF4 gamma chain involves a missense mutation, that is, a substitution of a guanine for an adenine base resulting in a concomitant change in the amino acid residue at position 415 from asparagine to aspartic acid. The position of the IF4 mutation can be correlated to variants of human lambda and kappa chains which occur in homologous regions. The position of residue 415 with respect to that of the second pseudosubunit half-cystine is very similar to that of the O₂ (in lambda)

a, MOPC-21 IgG/Kappa, indicating amino acid residue numbers involved in disulphide loops. IF4: 'missense' mutation at 415, no deletions. mRNA for IF4 and wild-type, normal, that is 17S. b, IF2: Deletion of C_{H1} region from position 120–215. Size of the mRNA is 16S, reduced from normal. Structure shown is as proposed by Secher *et al. (Immun. Rev.* **36**, 51; 1977). c, IF1: Deletion of C_{H3} (indicated by shaded area on diagram) due to 'nonsense' mutation at position 358. mRNA normal size. d, IF3: A two-base 'frameshift' mutation at position 341, results in an aberrant sequence (wavy line) to position *354* at which a premature 'ochre' termination codon occurs, deleting the rest of the C_{H3}. mRNA normal size.

and *Inv* (in kappa) markers in human light chains (Porter, *Defence and Recognition*, Butterworth, London, 1972).

The apparent preferred or pre-selected sites at which these spontaneous mutations are found and the supporting evidence from studies on human Ig chains sustain the hypothesis of the non-random nature of these mutations. However, there are some reservations. Allowances must be made for the relatively small number of mutants isolated and studied, and for those mutations which are lethal or those resulting in rapid degradation of the protein, because these would not have been detected by Milstein's procedures.

The results of Milstein and co-workers show that mutational events previously only shown to occur in prokaryotes and yeasts also occur in eukaryotes. Alternative explanations for their observations, such as post-transcriptional and/or post-translational modifications, are refuted or unnecessary as each mutation studied can be explained by simple classical genetic events. There is one cautionary note on account of recent reports on the post-transcriptional processing of the mRNAs coding for globin (see *News & Views* 264, 397, 9; 1977), oval-

bumin (Breathnach *et al.* *Nature* 270, 315; 1977), and immunoglobulin light chain (Wall & Gillmore-Herbert *Proc. natn. Acad. Sci. U.S.A.* 75, 342; 1978). It is suggested that looping, excision and repair are all steps involved in the processing of HnRNA to mRNA. Errors in one or more of these processing steps could well result in anomalous Ig chains.

One other interesting point is the observed mutational rate for the P3 cell line which is calculated at 10^{-5} per cell per generation (Adetugbo *et al.* *Nature*, 265, 299; 1977). This is two to three orders of magnitude higher than the rate detected for other mammalian cell lines.

These issues, namely the apparent preferential sites of mutation and their similarity to other reported studies and the high rate of mutation found for Ig gamma chains in this cell line, argue for the explanation that this cell line and possibly other Ig-producing cells experience a higher and more specific rate of spontaneous somatic mutations. Whether this mutational mechanism is the same as that which effects antibody diversity by mutations in the hypervariable residues in the V region of non-committed cells remains to be shown. □

units) sometimes in regular alignments, sometimes less ordered. Also visible was a network of cytoplasmic filaments (trabeculae) that underlies the membrane, possibly playing a role in aligning the receptors.

It has not been possible so far to identify visual counterparts of two other probable membrane components, the acetylcholine ionophore (or ion conductance modulator) and acetylcholinesterase. The former represents the channel through which cations are translocated after acetylcholine finds its receptor. In the laboratory of J.-P. Changeux (Pasteur Institute) the ionophore has now been tentatively identified as a 43,000 molecular weight protein component of *Torpedo* fragments that is distinct from the 40,000 molecular weight receptor. Confirmation is being sought by a combination of the techniques of reconstitution, cross-linking and immunocytochemistry.

Molecular information on acetylcholinesterase was reviewed by J. Massoulié (Ecole Normale Supérieure, Paris). A spectrum of molecular forms can be purified from *Torpedo*, the most complex and major form of which consists of three tetrameric clusters of subunits each of which is linked to one strand of a triple-stranded tail. The tail is collagen-like not only in being triple-stranded but also in its hydroxyproline and hydroxylysine content and its sensitivity to collagenase. In low ionic strength solutions, the tailed molecules aggregate apparently by their tails and under the influence of polysaccharides such as chondroitin sulphate. Massoulié proposed that it is through the chondroitin sulphate (-like) molecules of the basal laminar that acetylcholinesterase is correctly positioned to meet its substrate. Parallel investigations of the acetylcholinesterases of rat muscle have revealed a similar picture. One of the rat acetylcholinesterase forms is localised to the muscle endplate, disappears on denervation and is induced in cultured embryonic muscle cells only when they are cocultured with neurones (spinal cord explants). So neat is this picture, the more the pity that it turns out not also to be true of the muscle endplates of man, a species with an altogether more complicated set of acetylcholinesterase forms.

Returning to the acetylcholine receptor, S. Fuchs (Weizmann Institute) described two therapeutic approaches to myasthenia gravis. This disease results from auto-immunity to the receptor and can be studied in animals with experimental autoimmune myasthenia gravis

Neurobiology at Heidelberg

from Peter Newmark

THE attractions of acetylcholine-mediated neurotransmission as a model system for probing the molecular aspects of neurobiology are many. Not only is acetylcholine the least 'putative' of neurotransmitters but there already exists a wealth of information on its release from nerve terminals, its postsynaptic receptors and its inactivation by acetylcholinesterase. In considerable part this is due to the electric organ of fishes such as *Torpedo* spp, which possess an exceedingly rich cholinergic innervation. Without *Torpedo*, EMBO would surely not have had the nerve to invent the quasi-discipline of Molecular Neurobiology*.

Ultrastructural studies of the postsynaptic membranes of the electric organ of *Torpedo* have already revealed structures that can be tentatively identified as the acetylcholine receptors *in situ*. These consist of intramembranous particles on the cytoplasmic face of freeze-fractured membranes and a lattice-like structure on the surface of excitable microsacs derived from the organ. J. Heuser and S.

Salpeter (University of California, San Francisco) using the quick (2-ms) freeze technique that has been so successfully applied to the study of exocytosis at the presynaptic cholinergic terminals, have now been able to look for both intramembranous and surface particles in intact, unfixed *Torpedo* tissue. On several occasions they have managed to capture both simultaneously. On the true presynaptic surface the particles are densely packed in a fairly regular lattice; on the cytoplasmic freeze fracture face there are considerably fewer particles. To reconcile this numerical difference, Heuser proposed that all the particles (acetylcholine receptors) are transmembraneous but that freeze fracture tends to split them in half leaving no visible structure on the fractured faces. Only a minority of particles remain intact to form the intramembranous particles revealed by freeze fracture.

Further views of the presumed receptor on the true postsynaptic surface were obtained by freeze etching. A plan view of the surface showed dense arrays of receptor clusters (containing two, three or four

* The fourth EMBO annual symposium, held at Heidelberg, 3-7 May, 1978.

Peter Newmark is Biological Sciences Editor of Nature.

(EAMG) induced by immunisation with purified *Torpedo* receptor. If, however, the receptor is denatured before administration it does not result in EAMG and can delay, prevent or even reverse the EAMG produced by the native receptor. One possible explanation is that the antibody to the denatured receptor competes with that for the native form for binding to the muscle acetylcholine receptors of the experimental animals. The other therapeutic approach outlined by Fuchs is to try to supplement artificially the natural antibodies (anti-idiotypic antibodies) against the anti-acetylcholine receptor antibody with the intention of neutralising the harmful effects of the latter. Promising anti-idiotypic antibodies have been produced (see next week's *Nature* 273, 543; 1978) but not yet in high enough titres for their efficacy to be tested on EAMG.

Although the behaviour of acetylcholine conforms with that expected of a classical neurotransmitter, the neuropeptides move in more mysterious ways. What for example is substance P doing co-existing with 5-hydroxytryptamine in the neurones in the brain stem (T. Hokfelt, Karolinska Institute)? And where does that leave Dale's principle (although Hokfelt questioned its authenticity) of one neurotransmitter per neurone.

A suspicion, but no more than that, of a new brain peptide came out of the work of C. Schaller (European Molecular Biology Laboratory) on *Hydra*. She has partially purified several morphogenetic substances from *Hydra* including a head activator that seems to be a neuro-secreted peptide. Using the *Hydra* bioassay that was developed for the head activator Schaller has now discovered bioactivity in extracts of mammalian hypothalamus and intestine.

Oxytocin and vasopressin are peptide hormones that are synthesised in neurones of the supraoptic nucleus and paraventricular nucleus together with their individual carrier proteins, the neurophysins. M. Brownstein (National Institutes of Health) reported evidence that indicated the synthesis of both neurophysins in large precursor forms with sequential processing during their axonal transport to the posterior pituitary. The availability of Brattleboro rats that lack vasopressin and its associated neurophysin allowed the two neurophysins and their precursors to be distinguished. The way is now clear to determine whether, as is suspected, one precursor contains oxytocin and the other vasopressin, in addition to the respective neurophysin.

Even the more classical amine and amino acid neurotransmitters are looking considerably less classical these

Why subduction volcanism varies

from Peter J. Smith

REGIONS above the descending oceanic lithosphere are associated with high heat flow and active volcanism, as is well known. Conventional wisdom has it that these phenomena are caused by frictional heating in the slip zones of descending plates or by upwelling of hot mantle material induced by descending plates. But Ito (*J. geophys. Res.* 83, 262; 1978) claims that neither hypothesis can explain all the relevant petrological and geophysical observations. In island arcs, for example, volcanism results predominantly in andesites and (to a lesser extent) dacites, but in some cases also in basalts. So why, asks Ito, should some volcanic centres be involved in this 'bimodal' eruption?

To account for such puzzling phenomena, Ito follows Jischke (*J. geophys. Res.* 80, 4809; 1975) and others in appealing to a model in which there is, in the slip zone between a descending plate and the overlying asthenosphere, not only fluid produced by partial remelting of

subducted oceanic crust (in the lower regions) but also liquid drawn from the asthenosphere by suction (in the upper reaches). Where the new model differs from the old, however, is in assuming the density of the liquid-enriched fluid to be up to 10% lower than that of the asthenosphere and, more importantly, its viscosity to be much lower (perhaps as low as $10 \text{ g cm}^{-1} \text{ s}^{-1}$). The consequence of these crucial differences, as Ito shows at length mathematically, is an upward flow.

Support for the validity of Ito's basic assumption comes from seismic data obtained by Okada (DSc Thesis, Hokkaido University, 1977) who claims to have detected a thin low velocity layer along the upper edge of the plate descending beneath Japan. But be that as it may, Ito's model leads to an obvious explanation of the occurrence of both andesites and basalts in regions of the Earth's surface above descending plates. The former could result from partial remelting of subducted oceanic crust, as is commonly supposed, whereas the latter could be derived from the liquid drawn from the asthenosphere.

Peter J. Smith is a Reader in the Department of Earth Sciences at the Open University.

days. J. Glowinski (College de France) outlined evidence gleaned from the use of push-pull cannulae inserted into regions of the brains of anaesthetised cats, that clearly demonstrated the physiological and functional release of dopamine from the dendrites in the substantia nigra of neurones whose terminals released dopamine in the striatum. The dendritically-released dopamine functions as a neuromodulator through presynaptic receptors on a variety of neurones, including dopaminergic neurones themselves. Further experiments showed an intriguing reciprocal relationship between the nigrostriatal dopaminergic pathways of the two halves of the brain. For example, light flashes to the right eye resulted in increased release of dopamine from the ipsilateral striatal terminals with decreased release in the corresponding substantia nigra accompanied by just the opposite effects in the contralateral pathway. How this happens is not known but an involvement of the dendritically released dopamine is suspected.

Further complexity is introduced into the study of 'neurotransmitter' roles when account is taken of retrograde transport. That is, for example, the transport of glycine—a probable transmitter—from the nerve terminals in the tectum of the pigeon where it

is picked up, back to the cell bodies in the isthmi pars parvocellularis nucleus, (M. Cuénod, University of Zurich). Does it deliver a message? And equally does the quintessential neurobiological molecule, nerve growth factor (NGF), deliver a message to those parts of the neurone to which it is transported retrogradally along the smooth endoplasmic reticulum after internalisation at the nerve terminals (H. Thoenen, Biozentrum, Basel)? The answer for NGF is much more certainly yes than it is for glycine.

NGF is a specific growth factor for sympathetic neurones. A new model system for investigating its mode of action is a clonal cell line (PC12) derived from a transplantable rat adrenal medullary pheochromocytoma. In response to NGF, PC12 cells stop dividing, extend long neurites and become electrically excitable. P. Calissano (Harvard Medical School) presented tentative evidence for the internalisation of NGF by PC12 with the possibility that it reaches the nucleus. Calissano proposed that internalisation first allows NGF to interact with actin filaments—which it does *in vitro* to form stress fibre-like paracrystals—resulting in neurite extension. And then some NGF would reach the nucleus where it would interact with the tubulin—which it does *in vitro* to

form microtubules—of the spindle, stopping mitosis.

The branching of neurites from embryonic sensory or sympathetic ganglia under the influence of NGF *in vitro* has been studied by D. Bray (MRC Cell Biophysics Unit, London). A vectorial analysis of the branches made by growing neurites suggested that the growth cones are adhesive and thereby exert tension on the neurites which in turn are forced to seek maximum adhesivity. *In vivo* this mechanical tension could enable neurites to test their surroundings and select paths which are most adhesive. The kind of molecule involved in such adhesion is slowly yielding to the purification procedures of the laboratory of G. Edelman (Rockefeller University). His group has isolated a cell adhesion molecule (CAM) from retinal tissue culture that affects the disposition of neurites in histiotypic aggregates and in culture.

Finally the neurofilament—one of the more mysterious of neurobiological

structures—was discussed by D. Gilbert (MRC Cell Biophysics Unit, London). His own studies centre on the neurofilaments of the giant axon that runs down the back of the fanworm *Myxicola*. The axon is packed with neurofilaments lying in parallel and forming a long-pitch helix. Each neurofilament consists of seven subunits, again lying in parallel. And a subunit is composed of four similar α -helical protein chains each of about 150,000 daltons. The protein is vulnerable to proteolytic enzymes the action of which accounts for all the smaller proteins that appear as minor components of electrophoretic separation. Gilbert also raised the vexed question of the relationship of *Myxicola* neurofilaments to other neurofilaments/intermediate filaments. Proteolytic degradation may be the solution to the enigmatic variations of their molecular weights, but the functions of intermediate filaments are still not immediately clear (see *News & Views* 272, 577; 1978).

that in animals already studied.

Over the past two years much attention has been devoted to the effects of $24,25(\text{OH})_2\text{D}_3$ in man. J. Kanis (Churchill Hospital, Oxford) reported that this metabolite will increase calcium absorption—when measured by the retention of whole body calcium—in anephric patients and subjects with chronic renal failure. In addition, doses of 1–10 μg of $1,24,25(\text{OH})_2\text{D}_3$ did not produce this effect in the chronic renal function patients suggesting that the activity of $24,25(\text{OH})_2\text{D}_3$ is independent of the hydroxyl group in the 1 α position. Kanis tentatively proposed that $24,25(\text{OH})_2\text{D}_3$ may be a hormone involved with mineralisation of bone, even though the metabolite is without effect in the treatment of osteoporosis.

S. Edelstein (Weizmann Institute of Science, Israel) also believes $24,25(\text{OH})_2\text{D}_3$ to be involved in bone mineralisation. In chicks made vitamin D deficient and subsequently D replete by feeding $[1,2\text{-}^3\text{H}]$ - and $[4\text{-}^{14}\text{C}]$ -vitamin D_3 , Edelstein and co-workers observed little $1,25(\text{OH})_2\text{D}_3$ in bone. He says there is no $1,25$ receptor in bone and neither is there a vitamin D-dependent calcium-binding protein (CaBP). Edelstein suggests that workers who claim to have detected this CaBP in bone may not be using techniques which are sufficiently rigorous to remove contaminants. As for the foetuses of pregnant rats, Edelstein reports a large accumulation of $24,25(\text{OH})_2\text{D}_3$ in the bone compared with the intestine and other tissues. The newborn rat pups have high $25(\text{OH})_2\text{D}_3$ -24-hydroxylase activity, even in mothers kept on a low calcium diet.

In a preliminary experiment Edelstein's group found the bone morphology of chicks treated with $24,25(\text{OH})_2\text{D}_3$ to be similar to that of chicks treated with vitamin D_3 . This tends to confirm an earlier experiment of this group, in which it was found that the bone morphology of rachitic chicks treated with $1\alpha(\text{OH})\text{D}_3$ was still abnormal. Edelstein thus believes that $24,25(\text{OH})_2\text{D}_3$ is also required for normal bone growth.

Some advances in assay techniques were also reported at this meeting. C. M. Taylor (University of Manchester) reported that a nuclear receptor protein for $125(\text{OH})_2\text{D}_3$ —first reported by Lawson & Wilson (*Biochem. J.* 144, 573; 1974)—could be used in the competitive protein binding assay to measure this metabolite. J. L. H. O'Riordan (Middlesex Hospital, London) reported the development of a radioimmunoassay for $1,25(\text{OH})_2\text{D}_3$ using antisera against $1,25(\text{OH})_2\text{D}_3$ -25-

Resurgence of vitamin D

from Alastair Hay

TWENTY years ago many clinicians were doubtful about any major role for vitamin D in bone mineral homeostasis. If, however, the papers under discussion at a recent international workshop* were anything to go by, those doubts can no longer be sustained; vitamin D has begun to dominate the field. One of the most intriguing questions—still unanswered—is that of the function of vitamin D in fish. Generations of children have been the chief beneficiaries (or sufferers) of the discovery of the antirachitic properties of the lipid extracts of cod and halibut liver. But no-one has yet been able to demonstrate in a satisfactory manner how these fish came by their vitamin and why in some fish it is present in such high concentrations. Some interesting hypotheses have of course been advanced to explain the presence of this vitamin in fish. Zooplankton has been suggested as the source. Bills on the other hand thought that fish could have an enzyme for converting the precursor, 7-dehydrocholesterol, into vitamin D. D. R. Fraser (Dunn Nutrition Laboratory, Cambridge) is of the opinion that fish do have an enzyme to produce vitamin D but that the route is more complicated than the one initially postulated by Bills.

Fraser looked at the influence of light on the synthesis of the possible vitamin D precursor, 7-dehydrocholesterol. He has found 7-dehydrocholesterol in several fish tissues, with the highest concentrations in fish skin. *In vitro* experiments on sterol synthesis in fish skin using ^{14}C -mevalonic acid as the sterol precursor showed that when incubated in the dark, one of the products is 7-dehydrocholesterol. If the incubation is carried out under incandescent light however, the 5,8-epidioxide of 7-dehydrocholesterol is produced. This helps to explain puzzling observations on the influence of light on the metabolism of dietary vitamin D supplied to vitamin D-deficient rainbow trout. Fish kept under constant illumination appeared to be able to make the vitamin D metabolite, 25-hydroxy-vitamin D ($25(\text{OH})\text{D}$) from some source other than dietary vitamin D. Dietary vitamin D seems not to be efficiently conserved and contributes little to the concentration of $25(\text{OH})\text{D}$. Fraser proposes that the conversion from epidioxide to vitamin D could be achieved enzymatically.

So far, Fraser has found none of the other vitamin D metabolites in fish which have been found in birds and mammals. It therefore seems that vitamin D metabolism in fish might indeed turn out to be very different from

Held at Kiriath Anavim, Israel on 5–9 March.

Alastair Hay is a Research Fellow in the Department of Animal Physiology and Nutrition, University of Leeds.

hemisuccinate conjugated to bovine serum albumin. Although the two proteins are a useful addition to the techniques for measuring $1,25(\text{OH})_2\text{D}_3$, the problem still lies in the extraction of this metabolite from serum and its preparation for assay.

One of the problems in parathyroid hormone (PTH) estimations is measurement of the hormone in situations of hypoparathyroidism. J. M. Zanelli *et al.* (Kennedy Institute of Rheumatology, London) reported a cytochemical bioassay for PTH which is sensitive for low circulating level of the hormone. The assay involves the use of guinea-pig renal cortex as the target tissue and two sensitive dose-dependent responses, the NADH dependent re-oxidation pathway in the *pars recta* of proximal tubules and the glucose-6-phosphate dehydrogenase activity in the distal convoluted tubule.

One potentially valuable therapeutic tool for the future was discussed.

Paget's disease of bone has in recent years been treated with the diphosphonate ethone-1-hydroxy-1,1-diphosphonate (EHDP) and calcitonin with mixed success. So the report of O. L. M. Bijvoet (University Hospital, Leiden) that a new diphosphonic acid—amino-hydroxypropane diphosphonic acid (APD)—is effective in treating Paget's disease was welcome news to many clinicians. Bijvoet reported that oral administration of 6–30 μM of APD per kg body weight per day to patients, corrected hydroxyproline excretion within weeks and led to a reduction in alkaline phosphatase levels. This was accompanied by a positive calcium balance, and no osteomalacia was seen to occur. APD has yet to undergo further clinical trials but if the comment of one participant that "APD may be the penicillin of Paget's disease" is correct, then applied scientists have benefited once again from a bit of pure scientific research. $\square$

Are comets dirty snowballs or dust swarms?

from W. I. Axford and H. U. Keller

WITH regard to the article 'Are comets dirty snowballs or dust swarms' (*News & Views* 270, 558; 1978) we should like to discuss several important lines of argument and evidence that support the icy conglomerate nucleus model introduced by Whipple¹ and which seem to be completely inconsistent with the dust swarm model. During the past decade our knowledge of cometary physics has been considerably enlarged as the result of the use of modern techniques involving radio telescopes and ultraviolet spectrometers carried on rockets and satellites. Many of the results which have been obtained have confirmed predictions which were based on the icy nucleus model, whereas there are apparently no such predictions based on the dust swarm model and certainly none which has been similarly confirmed.

Poincaré has argued that physical theories "should establish connections between different experimental or observational facts, and above all, they should enable predictions to be made". Surely it is preferable to judge a theory on the basis of its successful predictions rather than appealing solely to the principle of Occam's razor. The latter is perhaps a useful means of judging the elegance and credibility of a model but cannot determine its validity. In

any case, as pointed out by Hughes, it does not give us grounds for believing that comets might be pure dust swarms in preference to icy conglomerates with their associated gas and dust envelopes.

The most important predictions of the icy conglomerate model are those which suggest that a comet is surrounded by a relatively dense, expanding gaseous coma comprised of sublimated ices (H_2O , CH_4 , NH_3 , CO , CO_2 , and so on, the so-called 'parent molecules') together with their dissociation products resulting from the interaction of these molecules with sunlight and the solar wind. Whipple required that typical cometary gas production rates be orders of magnitude greater than those determined from the available spectroscopic observations of the visible radicals in order to explain the 'non-gravitational' forces which apparently cause changes in the orbits of some comets. The appearance of the forbidden line of oxygen in bright cometary spectra led Biermann and Treffitz² to the conclusion that gas production rates similar to those required by Whipple must be prevalent. Similarly, Finson and Probst³ in their analysis of the features of the dust tail of comet Arend-Roland found that a gas production rate about two orders of magnitude larger than that derived from traditional observations of the visible cometary radicals was

necessary to enable dust freed from the nucleus to be accelerated in the dense inner regions of the expanding gaseous coma. Finally, Biermann⁴ predicted on the basis of the icy conglomerate model that a huge atomic hydrogen cloud must exist around comets as a result of the photodissociation of the parent molecules.

The first ultraviolet observations of comets by the OAO-2 satellite⁵ revealed the presence of very strong and extended Lyman- α emission (the resonance line of atomic hydrogen) around two comets, and so strikingly confirmed Biermann's predictions. The equivalent production rate of atomic hydrogen for a typical moderately bright comet at $r=1$ a.u. heliocentric distance was deduced to be more than 10^{29} s^{-1} (ref. 6), which is about two orders of magnitude greater than had been deduced solely on the basis of the visible radicals and in excellent agreement with values predicted by the authors mentioned above.

Water ice, or perhaps a water ice clathrate containing other compounds, has been favoured as the main constituent of the nucleus and is expected to provide the most abundant parent molecules^{7,8}. An investigation of the possible dissociation processes of possible parent molecules and in particular of H_2O suggested that a high velocity ($\sim 20 \text{ km s}^{-1}$) neutral hydrogen component should be present in the cometary coma⁹. The presence of this component was detected first on the basis of the distribution of Lyman- α intensity¹⁰ and confirmed recently by direct observations of the emission line profile by the Copernicus satellite¹¹. Furthermore, Sekanina and Miller¹², using the Finson-Probst approach, determined the dust and gas



A hundred years ago

WE are happy to state that a Commission appointed by the French Chamber of Deputies has reported favourably on the erection of a large observatory at Meudon, on the site of the Chateau which has been in ruins since the Franco-German War.

THE electric light display in the Paris streets and thoroughfares is becoming one of the attractions of Paris. Eight electric lamps have been placed in the Place de l'Opéra, twenty-four others in the Opera Avenue, and eight more on the Place du Théâtre Français.

From *Nature* 18, 6 June, 157 and 159; 1878.

W. I. Axford and H. U. Keller are at the Max-Planck-Institut für Aeronomie, Katlenburg-Lindau.

production rate of comet Bennett (1970II) from observations of its dust tail, and their results were found to be in excellent agreement with those derived from ultraviolet observations¹³. Again, in these cases the icy conglomerate model has given rise to predictions which have been later confirmed.

Such large production rates are seemingly incompatible with the notion that the gas is produced by desorption from dust grains. For example, if the solar wind is the source of the desorbed gas it is difficult to account quantitatively for the production rates of hydrogen¹⁴ and, more importantly, for the roughly equally strong production of OH (ref. 13), as well as for the existence of such parent molecules as HCN, CH₃CN, and H₂O which have been directly detected by radio techniques¹⁵⁻¹⁷. It should be noted in this respect that ice-coated grains cannot survive a close perihelion passage; indeed, even a single large nucleus must lose a layer of ice at least a metre thick due to sublimation during the perihelion passage with $r_{\min}=0.5$ a.u. The questions raised by these modern observations seem to us to constitute a serious, if not impossible, difficulty for the dust swarm hypothesis.

It is certainly not the case that the icy nucleus model is inconsistent with observations of the contraction of cometary comas as the comets approach the Sun. The apparent contraction is a very obvious consequence of the fact that the increase in gas production and in excitation of the coma of a comet approaching the Sun is counterbalanced by a decrease in the lifetime of the molecules and radicals due to dissociation and/or ionisation. For example, a very rough analysis indicates that, because the intensity of solar ultraviolet and of the solar wind flux both vary as $1/r^2$, the characteristic scale of the coma should vary approximately as r^n , where n is about 2 (ref. 18). The observations are consistent with such a variation¹⁹⁻²¹ and are not easily explained by the swarm hypothesis. It is also interesting to note that, although a dust swarm contracts in terms of its lateral dimensions, its length actually increases with decreasing heliocentric distance. However, as pointed out by Delsemme²¹, isophotes of monochromatic brightness profiles corresponding to neutral radicals are always nearly circular, barring rare cataclysmic events. It is difficult to understand how such results could be explained by desorption of gas from an elongated swarm.

The progress which has been made during the past decade has been associated with an increasing interest in the physics of comets and in particular their origin, composition and ultimate fate. Their possible significance as

Quasar luminosity

from F. Graham Smith

OBSERVATIONAL cosmology, although by now a very respectable subject, is still short of hard evidence about the way in which the Universe is evolving. A particular difficulty is that evolution affects the constituent bodies as well as the smoothed densities of matter and energy, so that we cannot easily test cosmological models by simple comparisons of distant and nearby galaxies. This is the main difficulty with tests involving counting populations of galaxies at different distances: a change in population at larger distances might be an indication of youth rather than a measure of the general expansion of the Universe.

Counts of distant quasars do in fact show very large evolutionary effects. A recent comparison between counts of bright and faint quasars (Green & Schmidt *Astrophys. J. Lett.* **220**, L1; 1978) shows a population increasing with distance, in all directions, as distance to the power 1.6. This could be due to an evolution in the average luminosity of quasars, or it could be due to a general cosmological evolution of density. In either case it is further evidence against any hypothesis that quasars are 'local' objects. There is, however, a serious problem in allowing for the evolution in luminosity if this observation is to be used for discriminating between closed and open models of the Universe.

A second test of cosmological models is obtained from a comparison of the redshift and the luminosity of quasars. This is particularly useful if measurements can be made at large distances, that is, at large redshifts. Unfortunately the luminosity is then very difficult to

measure directly, as the normally observed part of the spectrum is shifted out of the visible range. A very promising method for inferring luminosity from the intensity of emission lines has now been developed by Baldwin, Burke, Gaskell and Wampller (see page 431, this issue of *Nature*) which may remove the difficulty. They show that the luminosity of certain quasar emission lines increases as the third power of the continuum luminosity. They use the CIV line with rest wavelength 1,549 Å for redshifts between 1.1 and 2.5, and they show that the Mg II line at 2,800 Å can be used in the same way for quasars with lower redshifts. In this way, and with few assumptions, they can construct a 'Hubble diagram' for quasars covering most of the available range of redshifts.

The results favour a high value of the deceleration parameter q_0 , which excludes low-density models of the Universe. This agrees with previous but less reliable work on the relation between redshift and luminosity, and also with the evidence on populations. But there are still some important provisos; it is conceivable that the luminosity is itself affected by the local mean density of the Universe, or that it varies with cosmological epoch for some other reason. These provisos may be overcome eventually when the physics of quasars become better understood.

This new approach, in which the luminosity of a quasar can be measured from the intensity of emission lines, is important for the tentative support it gives for a closed Universe; it also shows that ground-based optical observations have a new and powerful way into critical cosmological questions.

F. Graham Smith is Director of the Royal Greenwich Observatory, Herstmonceux.

'building blocks' for the Solar System is now widely acknowledged and has stimulated efforts to perform *in situ* measurements from a space probe. A rendezvous mission, even if it involves a fast flyby, should be capable of providing most of the information we need to settle existing arguments and should also give us quite new insights into the nature of comets. Let us hope we do not have to wait too long for the realisation of such a mission, which is certainly feasible and possibly relatively inexpensive.

1. Whipple, F. *Astrophys. J.* **111**, 375 (1950); *ibid.* **113**, 464 (1951); *ibid.* **121**, 350 (1955).
2. Biermann, L. & Trefitz, E. *Z. f. Astrophys.* **59**, 1 (1954).
3. Finson, M. L. & Probst, R. F. *Astrophys. J.* **154**, 327 (1968); *ibid.* **154**, 353 (1968).
4. Biermann, L. *Jila Report* **93** (1968).

5. Code, A. P., Houck, T. E. & Lillie, C. F. *IAU Circ.* **2201** (1970).
6. Keller, H. U. *Space Sci. Rev.* **18**, 641 (1976).
7. Delsemme, A. H. & Swings, P. *Ann. Astrophys.* **15**, 1 (1952).
8. Delsemme, A. H. & Miller, D. C. *Planet. Space Sci.* **18**, 717 (1970).
9. Keller, H. U. *Mitt. Astr. Ges.* **30**, 143 (1971).
10. Keller, H. U. & Thomas, G. E. *Astron. Astrophys.* **39**, 7 (1975).
11. Drake, J. F. *et al. Astrophys. J.* (submitted).
12. Sekanina, Z. & Miller, F. D. *Science* **179**, 565 (1973).
13. Keller, H. U. & Lillie, C. F. *Astron. Astrophys.* **34**, 187 (1974).
14. Mendis, D. A. *Astrophys. Space Sci.* **20**, 165 (1973).
15. Huebner, W. F., Snyder, L. F. & Buhl, D. *Icarus* **23**, 580 (1974).
16. Ulrich, B. L. & Conklin, B. K. *Nature* **248**, 121 (1974).
17. Jackson, W. M., Clark, T. & Donn, B. *Proc. of IAU Coll. No. 25*, 272, Greenbelt, Maryland, 1974 (1976).
18. Mendis, D. A., Holzer, T. R. & Axford, W. I. *Astrophys. Space Sci.* **15**, 313 (1972).
19. Wurm, K. *Astrophys. J.* **34**, 312 (1939).
20. Mendis, D. A. & Ip, W.-H. *Astrophys. Space Sci.* **39**, 335 (1976).
21. Delsemme, A. H. *Proc. IAU Coll. No. 25*, 711, Greenbelt, Maryland, 1974 (1976).

articles

Classical instability of a naked singularity

F. de Felice

Theoretical Physics Institute, Department of Physics, University of Alberta, Edmonton, Canada T6G 2J1

The effective potential of the equatorial circular geodesics in Kerr metric show a sharp discontinuity from positive to negative values, when $a \rightarrow m$. It is shown that when a is in the small range $1.088 m > a > m$, stable circular orbits exist with negative energy (with respect to infinity); this implies that, if accretion takes place, a Kerr naked singularity is slowed down. The efficiency of this process increases as $a \rightarrow m$. Therefore, in the above range of a , Kerr naked singularities should decay to a black-hole state.

BLACK holes are generally accepted as real astronomical objects, but naked singularities are not. Although they are features of several vacuum solutions, their real occurrence in nature is a matter of controversy. The absence of any event horizon around them makes the pathologies of their geometry 'visible' from infinity¹. Unless we give up some well defined properties of space-time as the asymptotic predictability¹ or good causal behaviour², we have to assume cosmic censorship³ to ensure that naked singularities cannot develop from an initially non-singular state. This problem is far from being solved, however there are various indications^{1,3} which instil confidence in cosmic censorship as a physical law. Although I favour this view, good insights about physical processes which ultimately provide a cosmic censor mechanism can be deduced from the properties of the naked singularities themselves. Some attempts have been made, in the Reissner-Nordstrom metric⁴ and in the Kerr metric⁵, but the results remain debatable. It is generally believed that a naked singularity is unstable whenever it is surrounded by an ergoregion. In this case, processes can take place which involve particles with negative energy; these particles can only remain in the ergoregion and possibly fall into the singularity, damping out its energy as seen from infinity.

The Kerr metric with $a > m$ (a , specific angular momentum and m , mass of the metric source) is such an example and I show here that a process of this type mentioned causes a naked singularity to evolve towards a black-hole state. It turns out, in fact, that after accretion through an equatorial disk, a naked singularity (N.S.) progressively slows down, in contrast with the black-hole case.

The basic mechanism here is a process in which accreting particles moving in from infinity with positive energies and angular momenta, end up in the singularity with negative values of these parameters. This implies that while the black hole-N.S. transition is probably impossible^{6,7,12}, the reverse may be possible^{8,9}.

Effective potentials for circular orbits

The entire family of the equatorial circular geodesics in Kerr metric is described by⁹.

$$E_{\epsilon}(\rho, \alpha) = \rho^{-3/4}(\rho^{3/2} - 3\rho^{1/2} + \epsilon 2\alpha)^{-1/2}(\rho^{3/2} - 2\rho^{1/2} + \epsilon \alpha) \quad (1)$$

$$\epsilon = \pm 1$$

$$\lambda_{\epsilon}(\rho, \alpha) = \epsilon \rho^{-3/4}(\rho^{3/2} - 3\rho^{1/2} + \epsilon 2\alpha)^{-1/2}(\rho^2 - \epsilon 2\alpha\rho^{1/2} + \alpha^2) \quad (2)$$

in Boyer-Lindquist coordinates, with $\rho = r/m$, $\alpha = a/m$, $\alpha > 0$. We see later that equations (1) and (2) do not exhaust the entire family of the equatorial circular geodesics when $a > m$.

Here E_{ϵ} is the total energy per unit mass of the orbiting particle and λ_{ϵ} its angular momentum, again per unit mass and in units of m . In equations (1) and (2), $\epsilon = +1$ refers to co-revolving orbits and $\epsilon = -1$ to counter-revolving ones. Close to the metric source, only co-revolving orbits are allowed, so I shall take $\epsilon = +1$ and drop the sign label.

The energy $E(\rho, \alpha)$ diverges along the curve described by:

$$\alpha = \frac{1}{2}\rho^{1/2}(3-\rho) \quad (3)$$

The curve of equation (3) is shown in Fig. 1 and gives the innermost boundary (photon boundary) of the direct circular orbits. From Fig. 1, when $\alpha > 1$ (N.S. case), the photon boundary disappears; the only points of divergence for $E(\rho, \alpha)$ (as for $\lambda(\rho, \alpha)$) remain on the singularity itself at $\rho = 0$. Consequently, although in the black-hole case $E(\rho, \alpha)$ could never vanish in its domain of validity (the points of zero were all below the photon boundary), it now does so in a small range of α .

The locus where $E(\rho, \alpha)$ vanishes is described by the function

$$\alpha = \frac{(E)}{0} = \rho^{1/2}(2-\rho) \quad (4)$$

The graph of $\frac{(E)}{0}$ is shown in Fig. 1; the maximum occurs at $\rho = 2/3$ with $\alpha = 4\sqrt{2}/3\sqrt{3}$. It is clear that $E(\rho, \alpha)$ is negative

below the curve of equation (4).

To complete the study of $E(\rho, \alpha)$ we need to draw the locus of points in the (ρ, α) plane, where the energy is minimal.

This locus is given by

$$\alpha_{\epsilon}(\rho) = \frac{(E)}{m} = (1/3)\rho^{1/2}[4 + \epsilon\sqrt{(3\rho-2)}] \quad \epsilon = \pm 1 \quad (5)$$

which is shown in Fig. 1. The curve of Equation (5), without the branch $\alpha_{-}(\rho > 6)$ which describes counter-revolving

orbits only, identifies the direct innermost stable circular orbits (ρ_{ms}). The composite picture in Fig. 1 is a guide for drawing the general graph of $E(\rho, \alpha)$ when $\alpha < 1$ (Fig. 2a) and $\alpha > 1$ (Fig. 2b, solid line). When $\alpha > 1$, however, equations (1) and (2)

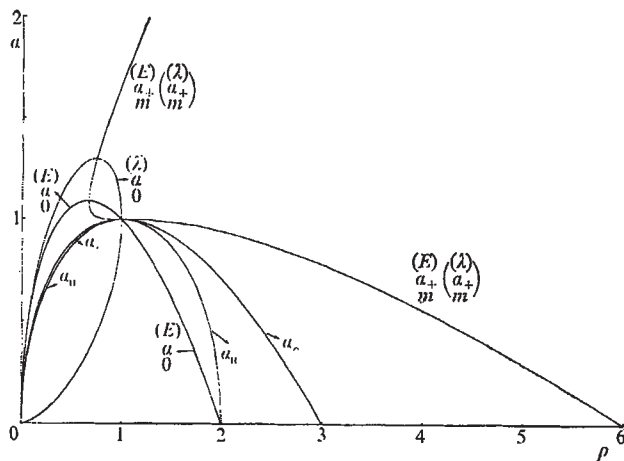


Fig. 1 Curve $\frac{(E)}{\alpha} \left(\frac{\lambda}{m^{\pm}} \right)$ defines the position of the innermost stable circular orbits (isco in the text) and is the locus where the function $E(\rho, \alpha(\rho, \alpha))$ is a minimum. Curve α/∞ defines the photon boundary and is the locus of points where $E(\rho, \alpha)$ diverges. Curve $\frac{(E)}{\alpha}$ is the locus of points where $E(\rho, \alpha)$ vanishes; Curve $\frac{(\lambda)}{\alpha}$ is the locus of point where $\lambda(\rho, \alpha)$ vanishes. Curve α_H defines the positions of the horizons when $\alpha \leq 1$.

do not represent all the physically possible circular orbits on the equator of the Kerr metric. In fact, the effective potential energy E being a solution of a quadratic equation, reads in general as:

$$\pm E(\rho, \lambda, \alpha) = f^{-1}[\alpha \lambda \pm (\alpha^2 \lambda^2 - f(2-\rho)(\lambda^2 - \lambda_0^2))^{1/2}] \quad (6)$$

where $f = \rho^3 + \alpha^2(2+\rho)$, $\lambda_0^2 = \rho(\rho^2 + \alpha^2 - 2)/(2-\rho)$. A simple investigation of equation (6) shows that (1) for a given value of λ we always have $+E > -E$; (2) $+E$ and $-E$ become zero at the same points when $\lambda^2 = \lambda_0^2$ and $\rho < 2$ (see ref. 10) but with distinct signs of the λ_0 s mainly $-|\lambda_0|$, for $+E$ and $|\lambda_0|$, for $-E$; (3) when λ is negative, the extremal values of $+E$ (the circular orbits described by equation (1)) are negative while the opposite is true for $-E$. Point (3) implies, therefore, that besides equation (1) when it becomes zero and negative with negative values of λ , one must also consider the corresponding function $-E(\rho, \alpha)$ which attains zero and positive values, but with λ positive, (Fig. 2b, dashed line).

The behaviour of the effective potential when $\alpha = 1$ needs consideration. In Fig. 1, the point $\rho = 1$, $\alpha = 1$ where all the curves intersect, is clearly a point of discontinuity for $E(\rho, \alpha)$ (as for $\lambda(\rho, \alpha)$). In fact, when $\alpha = 1$ equation (1) can be written as:

$$E(\rho, 1) = \frac{(\rho^{1/2} - 1)}{|\rho^{1/2} - 1|} (\rho + \rho^{1/2} - 1)(\rho^{1/2} + 2)^{-1/2} \rho^{-3/4} \quad (7)$$

$|\rho^{1/2} - 1|$ means the absolute value. In the limit $\rho \rightarrow 1$ we have:

$$L_{\rho \rightarrow 1+} E(\rho, 1) - L_{\rho \rightarrow 1-} E(\rho, 1) = 1/\sqrt{3}$$

this behaviour is shown in Fig. 2c. The step-like jump at $\rho = 1$ is only due to the particular nature of the Boyer and Lindquist coordinates; in this case, in fact, all the coordinate radii are projected into the coordinate value $\rho = 1$ (ref. 9). Here the existence of the horizon makes it meaningless to investigate further the curves $\pm E$ below it. The discontinuity in the energy (as in the angular momentum) at the ρ_{ms} when $\alpha \rightarrow 1_{\pm}$ suggests an astrophysical implication. When $1 < \alpha < \sim 1.088$, Figs 1 and 2 show that around the N.S. one can find stable circular orbits with ρ from $+\infty$ to ρ_{ms} and energies and angular momenta ranging continuously from positive to negative values (Fig.

2). One can also find stable circular orbits in the innermost part of the field with $\rho < \rho_{ms}$, but with an energy and angular momentum distribution which is disconnected from the previous one. All the other orbits are unstable. Hence, an accreting particle coming in from infinity moves along Fig. 2b (solid line) down to $\rho = \rho_{ms}$ radiating away more than 100% of its rest mass energy; from there, unless its energy and angular momentum undergo a drastic change to positive values, the particle, of mass μ , can only plunge into the N.S. To an observer at infinity this process leads to a decrease of the source energy of an amount $\mu E_{ms} = \{1 - 2/3 \rho_{ms}\}^{1/2} \mu$, which accounts for the radiation excess which might be detected at infinity.

The naked singularity decay

We are facing a kind of continued Penrose mechanism of energy extraction which makes the accretion onto a N.S. not only enormously more energetic than in the black-hole case, but also affects the metric source.

To understand the last point better, we consider how much angular momentum is transferred to the N.S. as a result of this process. The angular momentum of this particle is given by equation (2); direct orbits carry a positive angular momentum whenever $\alpha < 1$, but this is no longer true when $\alpha > 1$. In this case $\lambda(\rho, \alpha)$ vanishes along the curve described by the function:

$$\frac{(\lambda)}{\alpha} = \rho^{1/2} [1 + \varepsilon(1-\rho)^{1/2}] \quad \varepsilon = \pm 1 \quad (8)$$

The graph of equation (8) is shown in Fig. 1. Comparing equation (8) with $\frac{(E)}{\alpha}$, we find, as expected, that the entire

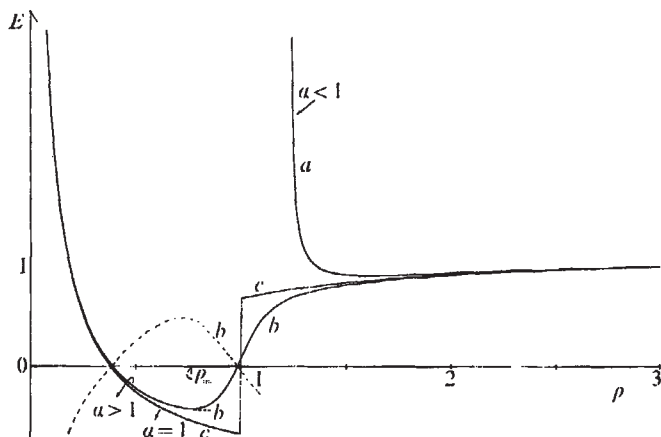
family of the circular orbits with negative energy lies in the region of the (λ, ρ) plane where $\lambda(\rho, \alpha)$ is negative. This readily implies that in the accretion process, the N.S. loses angular momentum as well.

To find the relative amount of the angular momentum loss with respect to the energy loss, we need the values of $\lambda(\rho, \alpha)$ at the ρ_{ms} . Differentiating equation (2) with respect to ρ we find that the extremes of $\lambda(\rho, \alpha)$ are given by the same function of equation (5) as for $E(\rho, \alpha)$ (ref. 10). In the range of interest, that is $2/3 < \rho < 1$ and $1 < \alpha < 4\sqrt{2/3}/\sqrt{3}$, we then have:

$$\lambda_{ms} = 2/3 [2\rho - 5/3 - (1/\sqrt{3})(\rho - 2/3)^{1/2}] [(1/\sqrt{3}) - (\rho - 2/3)^{1/2}]^{-1} \quad (10)$$

In the limit $\rho \rightarrow 1$, λ_{ms} reaches the value $-2/\sqrt{3}$. Accretion would slow down the N.S. with the same law as for when a

Fig. 2 The effective potential curves for circular orbits in the equatorial plane of Kerr metric. When $\alpha < 1$, the ρ_{ms} remain at $\rho > 1$ and $E_{ms} > 1/\sqrt{3}$; when $4\sqrt{2/3}/\sqrt{3} > \alpha > 1$, we find circular orbits with negative energy. When $\alpha = 1$ the effective potential curve undergoes a sharp discontinuity at $\rho = 1$.



blackhole is spun up^{11,12}. The evolution law reads:

$$m \frac{d\alpha}{dm_0} = \lambda_{ms} - 2\alpha E_{ms} \quad (11)$$

where m_0 is the rest mass that has gone into the source; from equations (1), (10), (5) and with ρ in the range $2/3 \leq \rho \leq 1$, equation (11) becomes:

$$m \frac{d\alpha}{dm_0} = -2/3 [(\rho_{ms} - 2/3)^{1/2} - (1/\sqrt{3})]^2 \leq 0 \quad (12)$$

In equation (12), the equality sign holds at $\rho_{ms} = 1$. The geometry parameter α decreases as matter accretes on the singularity until it attains the value $\alpha = 1$. Below this value the discontinuity at $\rho = 1$ ensures that λ_{ms} and E_{ms} become positive, reversing the sign of equation (12).

The slowing down, however, does not stop when $\alpha = 1$ but it continues, making the transition to a black hole irreversible. In fact, in the black hole the capture of the photons that are emitted by the accreting disk inhibits the spin-up of the hole to the maximal $\alpha = 1$. The hole capture cross section is greater for photons of negative angular momentum than for photons of positive. Therefore, the maximum value of the parameter α was shown to remain in the range $0.995 < \alpha < 0.999$ (ref. 12). Also for a N.S. which is affected by the photons which move strictly on the equatorial plane (the α used here = 2α used in ref. 10) the capture cross section for photons with negative angular momentum is greater than for photons with positive angular momentum, therefore their capture will make the slowing down more efficient. Not all the photons which are emitted from the equatorial plane escape to infinity; some, and especially those emitted in the inner part of the disk ($\rho \leq 1$), remain trapped around the N.S. This circumstance is particularly important because, as pointed out in refs 5 and 8, it accounts for the actual formation of the horizon at $\rho = 1$ when $\alpha = 1$.

Conclusions

The assumption that the space-time of a collapsing and highly rotating body is described by the Kerr metric with $\alpha > 1$ is

rather restrictive, as in general one would expect a considerable amount of gravitational radiation to be present. From the astrophysical point of view, however, a reasonable picture is that of a rotating body which collapses slowly and such that its space-time resembles Kerr metric with $\alpha \geq 1$. As it shrinks, it probably attains a disk-like form and so, most of the material which 'follows' the collapse will lie in the equatorial plane; an equatorial disk of accretion, therefore, seems likely to form.

The above calculations have shown that as $\alpha (> 1)$ is below the value 1.088, the inner part of the disk is made of negative energy particles. For this reason accretion takes place much more energetically than in the black-hole case; the efficiency for converting accreted mass into outgoing radiation is given by $\eta = 1 + E_{ms}$ and even greater than that if the photon contribution is considered.

In this case the N.S. slows down to the black-hole state; the proper time scale of this evolution clearly depends on the accretion model but this will be considered elsewhere.

When $\alpha > 1.088$ centrifugal forces prevail so as to prevent the peculiar circumstances of circular orbits with negative energy in the inner part of the disk; accretion will then lose efficiency.

I thank Professor W. Israel, and Dr A. Treves for helpful discussions. This work was partially supported by NRC Grant A-3030; financial support from the CNR of Italy is acknowledged.

Note added in proof: Recent calculations show that the ratio da/dm_0 in eqn (11) is negative for all $\alpha > 1$.

Received 19 December 1977; accepted 13 February 1978.

1. Hawking, S. W. *Commun. math. Phys.* **25**, 152 (1972).
2. Carter, B. in *Black-Holes* (Les Houches Summer School, 1972).
3. Penrose, R. *Il Nuovo Cimento* Ser. I—1, no. 252 (1969).
4. De Ruelle, N. & Damour, T. preprint Univ. Cambridge (1977).
5. Demianski, M. & de Felice, F. preprint, Univ. Warsaw (1977).
6. Wald, R. *Ann. Phys.* **83**, 548 (1974).
7. Penrose, R. *Proc. Texas, Symp. Rel. Astrophys.* (1973).
8. Calvani, M. & de Felice, F. preprint, Univ. Padova (1977).
9. Bardeen, J. M. Press, W. H. & Teukolsky, S. *Astrophys. J.* **178**, 347 (1972).
10. de Felice, F. *Il Nuovo Cimento* **57B**, 351 (1968).
11. Bardeen, J. M. *Nature* **226**, 64 (1970).
12. Thorne, K. *Astrophys. J.* **191**–507 (1974).
13. Jang, P. S. & Wald, R. M. *J. math. Phys.* **18**, 41 (1977).

Relative quasar luminosities determined from emission line strengths

Jack A. Baldwin

Institute of Astronomy, Cambridge, UK

William L. Burke, C. Martin Gaskell & E. J. Wampler

Lick Observatory, University of California,
Santa Cruz, California 95064

Observations of flat radio spectrum QSOs selected from complete samples confirm the strong correlation between continuum luminosity and emission line equivalent width. The r.m.s. scatter about a mean relationship is about 0.6 mag. The data indicate that the luminosity of QSO emission lines increase as the 1/3 power of the continuum luminosity. Unless the zero point of the relationship between emission line equivalent width and continuum luminosity depends on redshift, both the local hypothesis and zero pressure models of the Universe in which $q_0 \approx 0$ are ruled out by the data at about the 99% confidence level.

using the C(IV) λ 1,549 line for an inhomogeneous sample of quasars¹, and we have now verified its existence using a complete sample chosen solely on radio properties. We have also found a similar correlation for Mg (II) λ 2,800. These correlations provide a luminosity calibration for QSOs which has an intrinsic r.m.s. scatter of ± 0.6 mag. We give here a modified Hubble diagram for QSOs using this calibration procedure, and find that models with values of $q_0 \sim 1$ –2 provide the best fit to the data.

Our QSOs were taken from two complete samples, Schmidt's² flat spectrum sample and sample 5 from Wills and Lynds³. Samples of QSOs with flat radio spectra were chosen because it has been suggested that these show a particularly strong correlation between C(IV) equivalent width and continuum luminosity. It is important that the identifications of the optical objects in these samples are based strictly on positional coincidences between the optical and radio positions so that the

The correlation between the equivalent widths of QSO emission lines and the luminosity of the nearby continuum has been found

samples should not discriminate against faint, weak-lined quasars. In the two samples there are 84 identified QSO candidates with high frequency radio spectral indices flatter than -0.5 ; 38 with redshifts sufficiently high so that C(IV) can be observed from the ground, 38 with low redshifts, and eight with unknown redshifts. We originally intended only to observe all objects with $z > 1.1$ and all objects with unknown redshifts, but with the evidence that the correlation between line strength and continuum intensity also exists for Mg (II) we now intend to observe all the QSOs in these samples. We have now observed 87% of the QSOs in the samples with $1.1 < z < 1.5$ and 50% of the QSOs with $1.5 < z$ and we now present our interim results.

Observations

All observations were obtained at the Cassegrain focus of the 120-inch Shane telescope using the ITS system⁴. Photometric observations were made on clear nights of good seeing using entrance apertures 8 arc s in diameter. Stars from the list given by Stone⁵ were observed several times each night to calibrate the system. By comparing observations of the same standards obtained on different nights we estimate that on photometric nights the ITS system calibration was determined to an accuracy of 3% while individual QSO continuum magnitudes, which are also affected by the added uncertainty introduced by shot noise were generally accurate to $\pm 7\%$. Individual equivalent width determinations are accurate to about $\pm 20\%$. Objects that were observed on several nights have mean continuum determinations accurate to better than 5% and equivalent width determinations accurate to $\pm 10\%$. These

errors are much smaller than the observed scatter in the correlations. The observations are listed in Table 1. Observations obtained on nights that were not photometric were used only to improve the signal-to-noise of the spectra to estimate the equivalent width of the lines. Because we believe that at least one object, DA406, is variable, care was taken not to combine uncalibrated data with the data taken at a completely different time.

The equivalent widths of the emission lines were obtained by planimetry calibrated computer plots of spectra taken with either 7 or 15 Å resolution. Absorption features were ignored: the tabulated equivalent widths in every case are our estimate of the strength of the line had absorption not been present. In most of our data He(II) $\lambda 1,640$ was visible at low intensity on the red wing of C(IV) $\lambda 1,549$. The strength of He(II) was estimated by eye and the line was removed from the C(IV) profile. In all cases this correction was small as in our data $W_{\text{He(II)}} \ll W_{\text{C(IV)}}$. The continuum magnitude was estimated by fitting a straight line to the QSO continuum and measuring its estimated intensity at C(IV). This differs slightly from the procedure used by Baldwin¹ who measured the continuum at a rest wavelength of $\lambda 1,450$ Å, 100 Å to the blue of C(IV). The observed continuum intensities were converted to observed magnitudes through the relationship $AB_v = -2.5 \log f_v - 48.60$.

Results

To verify the correlation between emission line widths and continuum intensity, the QSOs were divided into three redshift groups. The continuum luminosity within each group was adjusted to partly compensate for the small redshift differences.

Table 1 Observed continuum magnitude and line strengths

Name	z	Equivalent width		Continuum Magnitude		No. of		Sample†	ϕ ‡
		W_o (C (IV))	W_o (Mg (II))	C (IV) $AB_{v \text{ obs}}$	Mg (II) m_v	$AB_{v \text{ obs}}$	observations		
0112-017	1.366	25	—	18.42	18.69	—	2	WL	0.148
0146+05 = OC+079	2.36	236	—	21.6:	21.04:	—	1	S	0.485
0147+09 = PHL1186	0.269	—	179	—	—	17.57	—	B	0.457
0155-10	0.616	—	30	—	—	17.47	19.04	B	0.139
0202+31	1.466	32	—	18.23	18.38	—	2	S	0.295
0234+28 = 4C28.07	1.21	50	—	18.91	19.36	—	2	S	0.261
0405-12	0.575	—	23	—	—	15.05	16.74	B	0.503
0457+02 = OF+097	2.38	39	—	19.48	18.92	—	1	WL	0.261
0748+33 = OI+380	1.92	26	—	17.85	17.60	—	1	S	0.377
0922+005 = OK+037	1.720	32	—	17.94	17.85	—	1	WL	0.404
0955+326 = 3C232	0.529	—	63	—	—	16.39	18.23	B	0.569
1004-018	1.214	212	—	20.8:	21.24:	—	1	WL	0.405
1010+35 = OL+318	1.40	143	—	20.40	20.62	—	3	S	0.388
1012+02	1.374	30	—	18.22	18.47	—	1	WL	0.257
1018+34 = OL+331	1.40	84	—	19.03	19.25	—	2	S	0.471
1100+772 = 3C249.1	0.311	—	34	—	—	15.69	18.52	W	0.289
1217+02	0.240	—	39	—	—	16.54	19.88	B	0.068
1226+023 = 3C273	0.158	—	9.7	—	—	13.25	17.41	B	0.054
1356+022	1.331	36	—	18.82	19.12	—	4	WL	0.194
1402-012	2.522	47	—	18.37	17.73	—	1	WL	0.565
1449-012	1.314	50	—	18.15	18.44	—	3	WL	0.439
1545+210 = 3C323.1	0.264	—	67	—	—	16.59	19.72	B	0.291
1602-00	1.633	35	—	18.13	18.12	—	2	WL	0.386
1611+34* = DA406	1.401	67	33	18.7:	18.9:	18.01	18.22	WL	—
1615+029	1.336	21	—	17.78	18.07	—	4	WL	0.209
1656+34	1.94	62	—	18.72	18.45	—	2	S	0.519
2134+004 = PHL61	1.930	30	—	17.44	17.18	—	5	WL	0.512
2223+21 = DA580	1.96	17	—	18.19	17.91	—	5	S	0.166
2243-03	1.348	76	39	18.10	18.38	17.56	17.84	S	0.606
2251+11	0.323	—	28	—	—	15.96	18.72	W	0.179
2251+24† = DA587	2.328	10	—	18.58	18.05	—	2	S	—
2254+024 = OY091.3	2.090	28	—	18.15	17.77	—	4	WL	0.369
2328+10 = 4C10.73	1.49	27	31	18.83	18.95	18.24	18.36	S	0.125
2332-017	1.185	43	—	18.30	18.78	—	3	WL	0.323

* DA406 is variable; the magnitudes and equivalent widths listed are mean values. ϕ was not computed for this object.

† There are many strong absorption features at the position of C (IV) $\lambda 1,549$ and it was not possible to obtain a reliable estimate of the emission-line strength. ϕ was not computed for DA587.

‡ The following were sources: WL, Wills and Lynds²; S, Schmidt³; B, Baldwin¹⁴; W, Wampler¹⁵.

§ The values of ϕ are from C (IV) data except for QSOs with only the preliminary Mg (II) calibrations.

We define a compensated magnitude by $m_v = AB_v - (5/2) \log(z^2/(1+z))$, which exactly compensates only if $q_0 = 1$. By not mixing QSOs of substantially different redshift we minimise the possibility that a poor choice in cosmological models will introduce a bias in the slope of the regression line connecting the magnitudes to the equivalent widths. We can also investigate the possibility that different relationships hold for different redshift intervals. The equivalent widths were reduced to the QSO frame by dividing by $(1+z)$. Figure 1 shows the C(IV) data for the three redshift groups and also the flat spectrum data from Baldwin's earlier list of QSOs. There is no significant difference in the least squares fit to the data of each individual group. From all the present data we find $\log L_{\text{cont.}} = -(1.66 \pm 0.17) \log W_{\text{C(IV)}} + \text{constant}$. This implies that $W_{\text{C(IV)}} \propto (L_{\text{cont.}})^{-2/3}$ and hence that the C(IV) luminosity is approximately proportional to the cube root of the continuum luminosity, a conclusion reached earlier by Baldwin¹. Repeating this analysis for $q_0 = 0$ produces a $\log L_{\text{cont.}} - \log W_{\text{C(IV)}}$ relation with greater scatter and a slightly steeper slope (-2.02), but does not alter any of our conclusions here.

While these data imply that for flat radio spectrum QSOs the mean relationship between the continuum luminosity and the equivalent width of C(IV) is the same for objects with different redshifts, the scatter in the diagrams is larger than can be accounted for by observational error alone. This is not surprising because: first, QSOs are known to be variable, and on time scales of a few years or less it is believed that only the continuum varies⁶⁻⁸. Such variations would move points in our diagrams along a line given by $\log L \propto \log(1/W_{\text{C(IV)}})$ rather than $\log L_{\text{cont.}} \propto 1.66 \log(1/W_{\text{C(IV)}})$ and would introduce scatter in the diagram. Second, no allowance has been made for interstellar extinction, which affects only the continuum. Third, although we have used monochromatic continuum magnitudes measured at a convenient wavelength, the actual physical mechanism governing the relationship between the strengths of the emission lines and the continuum may well assert itself through some broad-band luminosity (perhaps the bolometric luminosity or the total number of ionising photons) which is not perfectly correlated with the continuum luminosity

Fig. 1 The relationship between the equivalent width of C(IV) and the continuum magnitude at $\lambda 1,549$. +, Data for flat radio spectrum objects taken from Baldwin¹. Other points are all from two complete samples as discussed in the text. ●, QSOs with $1.1 \leq z \leq 1.4$; ○, objects with $z > 1.9$; and △ objects with $1.4 < z < 1.9$. The line represents the best least squares fit to all this data. There is no evidence for a different relationship for QSOs at different redshifts. The error bar is our estimate of the error of points with three or more observations. See Table 1.

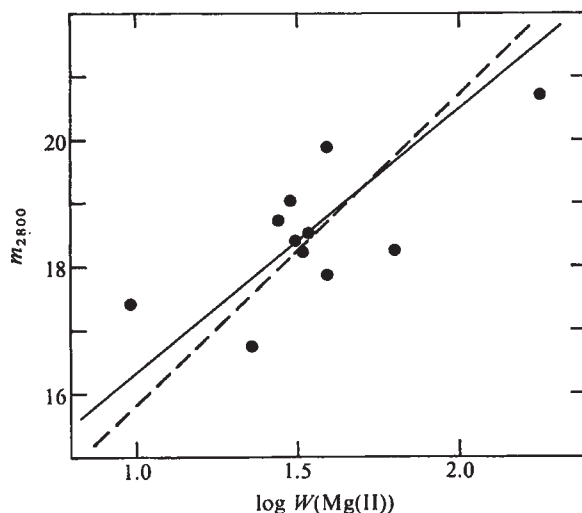
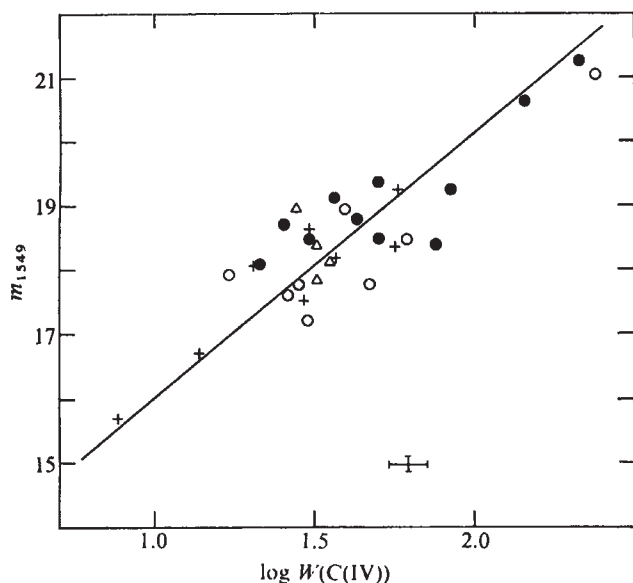


Fig. 2 The relationship between the equivalent width of Mg(II) and the continuum magnitude at $\lambda 2,800$. The dashed line is the least squares best fit to this data; the solid line is a line with the same slope as that found for the C(IV) data (Fig. 1).

at $1,549 \text{ \AA}$. This point cannot be properly examined until we have information on the continuum energy distributions over a much wider frequency range. Finally, differences in some other important physical parameter (age, for instance) could possibly produce scatter in the correlation.

We wish to use QSOs for cosmological tests because of their wide range of redshifts, but the C(IV) line can be observed from the Earth's surface only for redshifts greater than about $z = 1.1$. It is important to obtain data for lower-redshift QSOs. Davidson, *et al.*⁹ have observed 3C273 ($z = 0.158$) in the ultraviolet (UV) using a rocket-borne spectrometer, and they found that the spectrum is similar to that of high-redshift QSOs. Using the La luminosity calibration given by Baldwin they suggested a fit to cosmological models with $q_0 = +1$. We have examined the possibility of using ground-based observations of Mg(II) $\lambda 2,800$ to extend the luminosity calibration to redshifts lower than $z = 1.1$.

Like C(IV) $\lambda 1,549$, Mg(II) $\lambda 2,800$ is a resonance line which is relatively unaffected by blending with other nearby emission lines. However, the continuum near Mg(II) $\lambda 2,800$ may not be representative of the underlying 'power law' continuum of QSOs. The rocket UV data⁹ show a broad, unexplained rise in the continuum of 3C273 centred near $\lambda 2,800$. Baldwin¹⁰ also found that the continuum spectrum of many QSOs are enhanced in the region of Mg(II). Careful observations of low redshift QSOs will be necessary before we can firmly establish a luminosity calibration using Mg(II) $\lambda 2,800$. Nevertheless, preliminary results suggest that Mg(II) will be a useful indicator of continuum luminosity.

Mg(II) $\lambda 2,800$ equivalent widths and continuum intensities for a number of QSOs observed at Lick Observatory are given in Table 1. Figure 2 plots these data together with the least squares best fit straight line and a line representing the mean slope obtained from the C(IV) data after allowing a slight shift in zero point. As in the case of C(IV) $\lambda 1,549$ the continuum luminosities have been corrected by a factor $-2.5 \log_{10}(z^2/(1+z))$ to take into account differences in redshift.

Figure 2 shows that the slope of the regression line for Mg(II) agrees very closely with that found for C(IV). This observed agreement shows that, as for the C(IV) data alone, the slope is not strongly dependent on redshift. Either cosmological or effects intrinsic to the QSOs could introduce a shift in the zero-point of the curves found for the two redshift groups.

To extend the luminosity calibration to low redshifts using ground-based observations of Mg(II) it is necessary to determine the zero point of the Mg(II) calibration relative to C(IV)

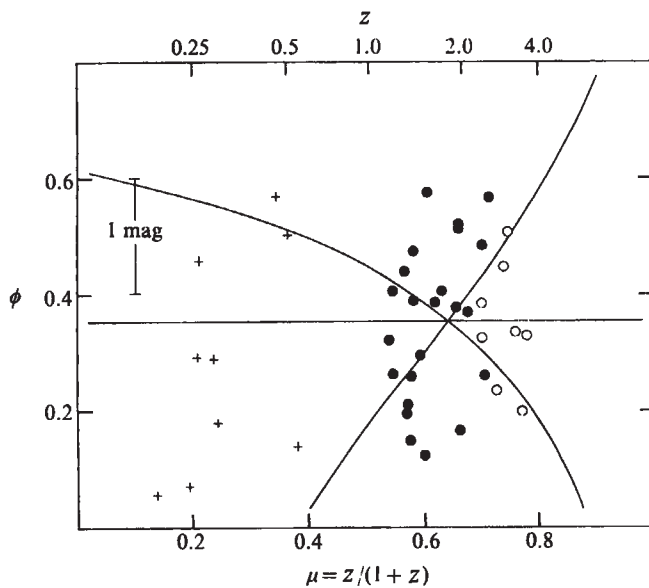


Fig. 3 The Hubble diagram for QSOs. The high redshift data ($\mu > 0.52$) is for flat spectrum objects taken from Table 1 (●) or from Baldwin¹ (○). +, Results from a preliminary Mg(II) calibration of low redshift QSOs. The horizontal line represents a $q_0=1$ cosmology; the line with negative slope represents $q_0=0$ cosmology, and the line with positive slope represents the local hypothesis (magnitude independent of redshift). Careful examination shows that the high redshift data alone is capable of rejecting the last two possibilities as lines with significant slope result in asymmetric distribution of residuals as a function of μ . The r.m.s. scatter of the data is about 0.6 mag—comparable with that seen in conventional Hubble diagrams of galaxies¹³. The preliminary low-redshift data (+) agrees with the high-redshift data and indicates the leverage to be expected from easily obtainable Mg(II) data.

by observing both lines in a number of QSOs. As a first attempt at this intercalibration we observed Mg(II) and C(IV) lines in 4C10.73 and PKS2243-03. We found that the zero point shift is near zero, the decrease in the strength of Mg(II) relative to C(IV) being approximately balanced by the increase in the continuum level. With parallel observations of only two QSOs this intercalibration is not very secure. Nevertheless, we believe that it is accurate to ± 0.5 mag. But our determination of q_0 given here rests only on the calibration of continuum luminosities using C(IV). The Mg(II) calibration data is given to show that it is in general agreement with the C(IV) calibration and to indicate that a much better determination will soon be possible.

Cosmological models

We have analysed these observations in the framework of homogeneous and isotropic cosmological models¹¹ extending the usual low-redshift analysis to the large redshifts encountered here. In these models there is a dimensionless distance coordinate χ related to metric distance by the scale-factor of the universe

$$d = R_0 \chi \quad (1)$$

the redshift z is related to time intervals

$$(t_0 - t) = z/H_0 \quad (2)$$

An arc-time η is defined so that the speed of light is unity in (χ, η) coordinates. Any cosmological model is determined by the evolution of the scale-factor, $R(\eta)$.

The propagation of light in a Robertson-Walker universe can be deduced from the adiabatic invariance of the photon number and the invariance of Planck's constant. A source of

intrinsic luminosity (spectral density) L_ν (W Hz^{-1}) will have an observed flux density F_ν ($\text{W cm}^{-2} \text{Hz}^{-1}$) given by

$$L_\nu/F_\nu = 4\pi R_0^2(1+z)S^2(\chi) \quad (3)$$

where ν' is the redshifted frequency

$$\nu = \nu'(1+z) \quad (4)$$

R_0 is the scale factor of the universe at the time of observation and $S(\chi)$ is either $\sin \chi$, χ , or $\sinh \chi$ for closed, critical, or open models respectively. Although χ is not observable, it can be related to the redshift z by the specific cosmological model through the function $R(\eta)$,

$$(1+z) = R(\eta_0)/R(\eta_0 - \chi) \quad (5)$$

where η_0 is the arc-time of observation. For example, in a closed, zero-pressure model one has

$$(1+z) = (1 - \cos \eta_0)/(1 - \cos(\eta_0 - \chi)) \quad (6)$$

The conventional Hubble diagram is a plot of the observed flux against redshift. Because the scale factor R_0 and the intrinsic luminosity L_ν are poorly known there is an unknown multiplicative factor in the flux density, hence we plot the logarithm of the flux density. While it is usual to plot this against $\log z$, it would be better to absorb both the singularity at $z = 0$ due to the inverse square law, and the singularity at $z = \infty$ due to the redshift into the measured flux. The quantity

$$\phi = (1/2)\log_{10}(z^2 F_\nu/(1+z)) \quad (7)$$

is finite from $z = 0$ to $z = \infty$, and isolates the geometrical effects peculiar to a particular space-time. Because ϕ goes to a finite limit as $z \rightarrow \infty$, the curves flatten out and for a large z contain little information. We remedy this by using instead of z the equally natural variable μ .

$$\mu = z/(1+z) = (\lambda_{\text{rec}} \cdots \lambda_{\text{emit}})/\lambda_{\text{rec}} \quad (8)$$

which ranges from 0 to 1 over the observable universe. In such a plot, models of different present ages η_0 are represented by curves of different slopes, rather than by curves of different curvature as in a conventional Hubble diagram. These theoretical curves are plotted using equation (5) for $\chi(z)$ and

$$\phi = \log(\mu/S(\chi)) + \text{constant} \quad (9)$$

Because our observations are at large redshifts, the deceleration parameter (q_0) is found by fitting a model, and determining q_0 from the age parameter η_0 . For zero-pressure models we have

$$q_0 = (1 + \cos \eta_0)^{-1} \quad (10)$$

Thus q_0 is related to the slope of the curves at $z = 0$ by

$$\left. \frac{d\phi}{d\mu} \right|_{z=0} = (\log_{10} e)(q_0 - 1)/2 \quad (11)$$

Estimating q_0

Figure 3 shows a plot of $\phi = 1/2 \log z^2/(1+z) - AB_{\nu}/5 + 0.832 \log_{10} W + 2.719$ against the redshift variable μ . The offset for ϕ_0 was arbitrarily chosen so that a 17.00 mag QSO at $z = 1$ with $W_{\text{C(IV)}} = 10$ has $\phi = 0$. Lines corresponding to various cosmological models are shown. The least squares best fit to the data with $z > 1.1$ gives a slope of 0.3 ± 0.3 . This value corresponds to $q_0 \approx 2$ for a 'dust' (zero pressure) universe. The uncertainties will be much reduced when the low redshift data from the Mg(II) $\lambda 2,800$ calibration can be used. The future ability to tie the Mg(II) calibrations to those of C(IV) will increase the baseline in μ by a factor of 3, and this, together with a reasonable increase in the number of data points should reduce the present slope uncertainty of ± 0.3 by a factor of about 6.

This data is unusual in three respects, however. First, there is no null hypothesis here; the horizontal line is in fact $q_0 = 1$.

Second, the scatter in the points is population scatter, not primarily measurement error. More data will not make the cloud of points contract, but will define its trend better. Third, the scatter in this plot is all in the vertical direction. This anisotropy allows us to fit a nearly horizontal line to what appears to be a round cloud of points.

Using only the C(IV) data we can draw several interesting conclusions. Were the continuum magnitude uncorrelated with redshift, as in the local hypothesis, then the residuals would have a systematic slope that would occur by chance in normally distributed data only 1% of the time. Another hypothesis is the low density, standard model of the universe. The residuals in a fit to a zero-pressure $q_0 = 0.03$ model again show a systematic trend in redshift; it would arise by chance only 1% of the time. A low-density universe with QSOs whose average properties do not depend on time is thus excluded. The reasonable fits to zero-pressure models correspond to a range of arc-time of $90^\circ < \eta_0 < 150^\circ$, over which q_0 is rapidly increasing, becoming infinite at $\eta_0 = 180^\circ$.

The formal value $q_0 \approx 2$ found here is, within the errors, consistent with the value $q_0 = +1$ found by Davidsen *et al.*, using Baldwin's¹ $\text{Ly}\alpha$ luminosity calibration, with the value $q_0 = +1.38$ found by Fang *et al.*¹² or with the value $q_0 = +1.6$ found by Kristian, Sandage and Westphal¹³ using galaxies with no correction for luminosity evolution. All these models require the Universe to be filled with as yet undiscovered material such as matter or radiation, or material described by a cosmological constant. These data do not distinguish between these options. If q_0 is as large as $+2$ the implied age of the Universe, using a dust model with $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is $9 \times 10^9 \text{ yr}$. This is uncomfortably short as modern lower estimates of the ages of the oldest stars are about 10^{10} yr and age determinations from the relative abundances of radioactive species give about the same lower estimate to the age of the Universe.

If the high value for q_0 obtained here is due to evolution of QSOs only the zero point and not the slope of the relationship between equivalent width and continuum luminosity can change. This is because the distribution of luminosities in our sample is approximately the same for the different redshift intervals and statistically permissible changes in the slope

would not substantially change our conclusion that neither the local hypothesis nor the $q_0 = 0$ curve is a good fit to the data. The large luminosities of QSOs make it highly unlikely that any one QSO could remain active for more than a small fraction of the age of the Universe. Thus at any redshift we expect to see QSOs of all ages; young newly-formed objects will coexist with QSOs near the end of their active phase. This is a completely different situation to that encountered in the study of galaxies; galaxies seen at low redshifts are believed to be evolved versions of those seen at high redshifts.

The agreement between our determination of q_0 from QSOs and the value obtained by Kristian, Sandage and Westphal for galaxies, implies that if evolution is important, the epoch dependent luminosity changes are similar for both QSOs and galaxies, which is curious in view of the different natures of the evolutionary process. These circumstantial arguments do not prove that there is little luminosity evolution; that must wait for more complete understanding of QSO physics. Nevertheless, the correlation between continuum luminosity and equivalent width of QSO emission lines provides an important constraint on models of QSO emission line regions and at least suggests that QSOs can be used as cosmological test particles.

We thank Howard French for assistance in the preparation of finding charts and help with observations.

Received 26 February; accepted 13 March 1978.

1. Baldwin, J. A. *Astrophys. J.* **214**, 679–684 (1977).
2. Schmidt, M. *Astrophys. J.* **217**, 358–361 (1977).
3. Wills, D. & Lynds, R. *Astrophys. J. Suppl.* (in the press).
4. Miller, J. S., Robinson, L. B. & Wampler, E. J. *Adv. Electronic Electron Phys.* **40B**, 693–698 (1976).
5. Stone, R. P. S. *Astrophys. J.* **193**, 135–137 (1974).
6. Sandage, A. R. *IAU Circ. No.* 1961 (1966).
7. Oke, J. B. *Astrophys. J.* **147**, 901–907 (1967).
8. Wampler, E. J. *Astrophys. J.* **148**, L101–L106 (1967).
9. Davidsen, A. F., Hartig, G. F. & Fastie, W. G. *Nature* **269**, 203–206 (1977).
10. Baldwin, J. A. *Astrophys. J.* **201**, 26–44 (1975).
11. Misner, C. W., Thorne, K. S. & Wheeler, J. A. *Gravitation* (W. H. Freeman, San Francisco, 1973).
12. Fang, L. Z., You-Yuan, Z., Fu-Zhen, C. & Yao-Chuan, C. *Acta ast. sinica* **17**, 134–146 (1976).
13. Kristian, J., Sandage, A. R. & Westphal, J. A. *Astrophys. J.* **222**, 383 (1978).
14. Baldwin, J. A. *Mon. Not. R. astr. Soc.* **178**, 67p–74p (1977).
15. Wampler, E. J. *Astrophys. J.* **153**, 19–26 (1968).

Pb, Nd and Sr isotopes in oceanic ferromanganese deposits and ocean floor basalts

R. K. O'Nions, S. R. Carter, R. S. Cohen, N. M. Evensen & P. J. Hamilton

Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964

The Pb-, Nd-, and Sr-isotope compositions of oceanic ferromanganese deposits, together with the Nd-, and Sr-isotope compositions of altered ocean-floor basalts are reported. These data are used to evaluate the sources of these metals in both the oceans and ferromanganese deposits and the extent to which ocean-floor basalts are a source of, or a sink for these metals.

CONCENTRATION of metals occur on the sea floor in two principal types of ferromanganese deposits. One of these, sometimes referred to as hydrogenous deposits¹, comprises the Mn-nodules and is widespread in areas of low sedimentation rate such as the abyssal parts of the ocean basins. The second type, referred to as hydrothermal deposits¹, is the metalliferous deposits associated with the tectonically active regions near

spreading centres in the oceans and is best developed on the East Pacific Rise.

There are various viewpoints concerning the source of metals in authigenic ferromanganese deposits. For example, Bostrom and Peterson^{2,3} have attributed the enrichment of Fe and Mn in East Pacific Rise sediments to local volcanic exhalatives, and Corliss⁴ developed a model whereby Fe, Mn, Co, rare-earths and other metals are liberated during the hydrothermal alteration of basalt and provide a source for these metals in metalliferous sediments. This contention was supported by high Mn-accumulation rates in these environments⁵ and by the isotopic composition of Pb in metalliferous sediments which is indistinguishable from that in mid-ocean ridge basalts⁶. However, the isotopic composition of U and Sr, and the relative abundances of the rare-earth elements in metalliferous sediments indicate derivation of these metals directly from seawater^{1,6}. It is clear that the hydrothermal alteration of oceanic basalts makes a small contribution to the

total Sr budget of the oceans^{7,8}, and seawater Sr is mostly derived from weathering of the continental crust. In contrast, the source(s) of the rare-earth elements in seawater is uncertain. The source(s) of metals in Mn-nodules has been similarly obscure. The Mn-accumulation rate in pelagic sediments and Mn-nodules averages about $1 \mu\text{g cm}^{-2}\text{yr}^{-1}$ (ref. 9) compared with about $35 \mu\text{g cm}^{-2}\text{yr}^{-1}$ in East Pacific Rise sediments⁵. Although the lower values for the Mn-nodules are consistent with derivation of all the Mn from mid-ocean ridges, Turekian¹⁰ has pointed out that sufficient Mn may be supplied to the oceans from estuaries to account for the average accumulation rate. The relative rare-earth element abundances in Mn-nodules differ from those in seawater and metalliferous sediments, particularly with respect to Ce, but no consensus has emerged as to the significance of this observation¹¹⁻¹³. Pb-isotope compositions of Mn-nodules^{14,15} measured so far are variable and seem to include lead from both mantle and continental sources.

The problem of the sources of metals in ferromanganese deposits in the oceans reduces to one of identifying and quantifying the fluxes of metals into the oceans from ocean basalt (= mantle) and continental sources. The degree of mixing between metals supplied from these sources depends upon their residence times in the oceans compared with oceanic mixing rates; the extent to which metals in ferromanganese deposits are provenance dependent will similarly reflect these parameters.

The fact that the continental crust has developed with a considerable fractionation of Rb/Sr, Sm/Nd and U/Pb relative to the mantle, coupled with large differences in oceanic residence times of Pb (10^6 to 10^8 yr)¹⁶⁻²¹, Nd (probably similar to Pb) and Sr ($\sim 10^6$ yr; similar to Ca)²², makes it possible to place constraints on the origin of ferromanganese deposits by a consideration of Nd-, Pb- and Sr-isotope compositions. Hence measurements have been made of $^{143}\text{Nd}/^{144}\text{Nd}$, $^{208}\text{Pb}/^{204}\text{Pb}$, $^{207}\text{Pb}/^{204}\text{Pb}$, $^{206}\text{Pb}/^{204}\text{Pb}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in metalliferous sediments from the East Pacific Rise and Mn-nodules from the Pacific Ocean and south-east Indian Ocean. In addition, $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios have been measured in drilled basalts from DSDP Leg 38 and dredged basalts from the south-east Indian Ridge.

The analytical techniques used in this laboratory for Nd and Sr isotope analyses have been described previously²³. Pb was separated by an electrodeposition technique similar to that described by Arden and Gale²⁴ and was isotopically analysed on a V.G. MM-30 mass spectrometer using a standard silica gel-phosphoric acid technique.

Nd, Pb and Sr in ferromanganese deposits

The Pb-, Nd- and Sr-isotope compositions have been measured in four manganese nodules and two samples of metalliferous sediment (Table 1). Three of the manganese nodule samples (A-50D; D-3D and B-15P) are from the North Pacific and the fourth (F-55) is from the South Australia Basin in the south-east Indian Ocean. The two samples of metalliferous

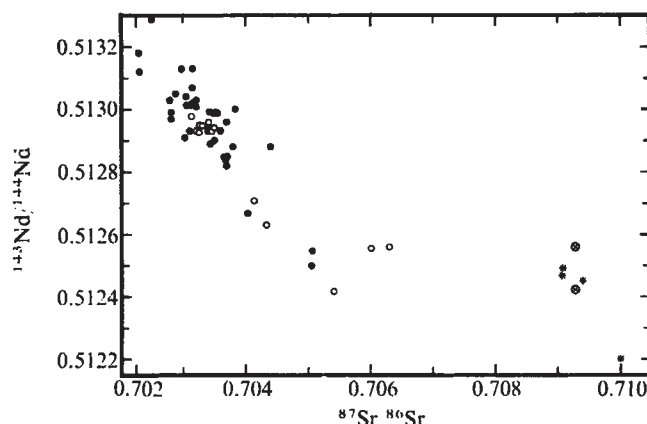


Fig. 1 Comparison of $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in unaltered oceanic (●) and continental basalts (○)^{23,28-30} with those in Mn-nodules* and metalliferous sediments (⊗) (see Table 1). Note that mid-ocean ridge basalts have $^{143}\text{Nd}/^{144}\text{Nd} \geq 0.5130$.

sediment from the East Pacific Rise were collected during cruise 19 of the RV Vema and are essentially identical to those described by Bender *et al.*⁵.

With the exception of F-55, the $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of the samples (Table 1) are similar to the reported values for seawater of ~ 0.7091 (refs 25-27). Sample F-55 from the South Australia Basin has an $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.71000 ± 4 , which is significantly higher than seawater; however, the acetic acid leachate of this sample is 0.70924 ± 5 . This suggests that non-authigenic continent-derived material with $^{87}\text{Sr}/^{86}\text{Sr} > 0.7100$ is an important component of F-55. The three manganese nodule samples from the North Pacific (A-50D, D-3D, B-15P) have indistinguishable $^{143}\text{Nd}/^{144}\text{Nd}$ values and average 0.51247 , whereas F-55 has a significantly lower value of 0.51220 ± 2 , consistent with the presence of continent-derived detritus with $^{143}\text{Nd}/^{144}\text{Nd} < 0.5122$. The metalliferous sediment samples (Table 1) exhibit a small range of $^{143}\text{Nd}/^{144}\text{Nd}$ ratios from 0.51242 ± 3 to 0.51256 ± 3 , similar to those of the North Pacific Mn-nodules. Collectively, these samples suggest that seawater has a $^{143}\text{Nd}/^{144}\text{Nd}$ ratio between 0.51242 and 0.51256 . Note that authigenic material replacing fish debris³⁹ has yielded a $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of 0.51215 , which may indicate that the Nd was only partly derived from seawater. As far as the rare-earth elements are concerned these results imply that the positive Ce anomaly in Mn-nodules, for example^{11,12}, and the negative Ce-anomaly in metalliferous sediments⁵ are not due to different sources for the REE. Rather, the rare-earth elements in both deposits are derived from seawater and the anomaly is a function of their mechanisms of removal from seawater.

The Nd- and Sr-isotope compositions of the Mn-nodules and metalliferous sediments are compared with those of recent unaltered basalts^{23,28-30} in Fig. 1. R.K.O., P.J.H. and N.M.E.²³

Table 1 Pb, Sr and Nd isotopic data for ferromanganese samples

Sample	$^{208}\text{Pb}/^{204}\text{Pb}^*$	$^{207}\text{Pb}/^{204}\text{Pb}^*$	$^{206}\text{Pb}/^{204}\text{Pb}^*$	$^{87}\text{Sr}/^{86}\text{Sr}^\dagger$	$^{143}\text{Nd}/^{144}\text{Nd}^\ddagger$
Mn-nodules					
A-50D ($32^\circ 42.0' \text{ N}; 158^\circ 14.5' \text{ E}$)	38.716 ± 84	15.627 ± 24	18.598 ± 17	0.70941 ± 4	0.51245 ± 3
D-3D ($9^\circ 07.0' \text{ N}; 105^\circ 00.0' \text{ W}$)	38.041 ± 52	15.514 ± 16	18.452 ± 12	0.70908 ± 5	0.51249 ± 3
B-15P ($11^\circ 40.7' \text{ N}; 109^\circ 45.5' \text{ W}$)	38.442 ± 15	15.560 ± 5	18.666 ± 6	0.70906 ± 5	0.51247 ± 2
F-55-16-1 ($37^\circ 09.7' \text{ S}; 121^\circ 13.4' \text{ E}$)	38.820 ± 50	15.638 ± 17	18.812 ± 15	0.71000 ± 4 (0.70924 ± 5)§	0.51220 ± 2
Metalliferous sediments					
V-19-53(40) (East Pacific Rise, 17° S)	37.929 ± 38	15.478 ± 13	18.494 ± 13	0.70928 ± 4	0.51256 ± 3
V-19-53(85) (East Pacific Rise, 17° S)	37.951 ± 25	15.486 ± 9	18.486 ± 10	0.70926 ± 5	0.51242 ± 3

*Errors quoted on uncorrected measured ratios are 2σ for within run statistics.

Data obtained on NBS-981 indicates a fractionation of $0.6 \pm 0.2\text{‰}$ a.m.u.⁻¹.

†Eimer and Amend SrCO_3 , $^{87}\text{Sr}/^{86}\text{Sr} = 0.70808$.

‡Normalised to $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$.

§Acetic acid leachate.

and DePaolo and Wasserburg²⁹ have shown that the $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of unaltered basalts from both oceanic and continental areas show a strong negative correlation. The Nd- and Sr-isotope compositions of the authigenic ferromanganese deposits plot to the right of the trend defined by these recent unaltered basalts and have distinctive isotope compositions. These distinctive compositions are imparted by a major contribution of radiogenic Sr and unradiogenic Nd from the continents to seawater.

Lead isotope compositions are variable in the ferromanganese samples (Table 1; Fig. 2) and cover a similar range in Pb-isotope composition to those reported previously by Reynolds and Dasch¹⁵. The $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ ratios of ocean-floor basalts^{31,32} are plotted in Fig. 2 for comparison. Figure 2 also shows two recent estimates of the composition of average modern terrestrial lead^{33,34}. The Mn-nodules have Pb-isotope ratios which include a substantial part of the variation in ocean-floor basalts and the estimated compositions of average terrestrial Pb, whereas the metalliferous sediments have ratios which are clearly in the field of ocean-floor basalts. The range of Pb-isotope compositions in Mn-nodules and metalliferous sediments can simply be explained as a mixture of Pb derived from ocean-floor basalts and the continents as originally suggested by Dasch *et al.*³⁵. The proportion of Pb from each source is difficult to estimate in a particular situation because of the considerable variation which exists in ocean-floor basalts and the uncertainty in the composition of continent-derived Pb. Nevertheless, the Pb-isotope data do demonstrate the high degree to which the Pb-isotope compositions of Mn-nodules and metalliferous sediments are provenance dependent, and the large component of Pb derived from ocean basalts in metalliferous sediments.

Nd and Sr in altered basalts

Previous studies of alteration in ocean-floor basalts³⁶⁻³⁷ have clearly demonstrated their susceptibility to contamination by seawater Sr. Measurements of $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in some DSDP Leg 37 basalts^{23,37} suggest that modification of $^{87}\text{Sr}/^{86}\text{Sr}$ ratios by seawater interaction may not be accompanied by significant changes in $^{143}\text{Nd}/^{144}\text{Nd}$ ratios. This

Fig. 2 *a*, $^{207}\text{Pb}/^{204}\text{Pb}$ against $^{206}\text{Pb}/^{204}\text{Pb}$ ratios in mid-ocean ridge basalts reported by Church and Tatsumoto³¹ (●), DSDP Leg 37 basalts reported by Cumming³² (▲), compared with estimates of average terrestrial lead(+) by Stacey and Kramers³³ (A) and Cumming and Richards³⁴ (B). *b*, $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ ratios in Mn-nodules (●) (North Pacific and south-east Indian Ocean) and metalliferous sediments (○) (East Pacific Rise).

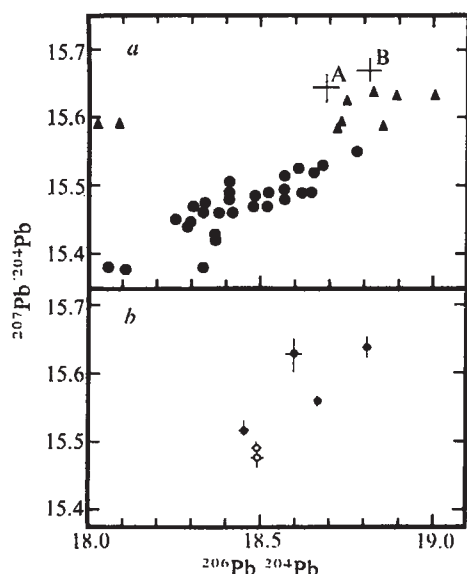


Table 2 Sr and Nd isotopic data for altered basalts

Sample	$^{87}\text{Sr}/^{86}\text{Sr}^*$	$^{143}\text{Nd}/^{144}\text{Nd}^\dagger$
DSDP Leg 38		
350-16-2	0.70400 ± 4	0.51307 ± 4
344-34-2	0.70388 ± 4	0.51306 ± 4
338-45-2	0.70352 ± 4	0.51302 ± 3
348-32	0.70439 ± 4	0.51316 ± 3
South-east Indian Ridge		
D4-A1 137° E.	0.70297 ± 6	0.51301 ± 3
D2-A4 134° E.	0.70575 ± 4	0.51308 ± 2
D8-B4 123° E.	0.70612 ± 5	0.51288 ± 3
D-11 B2 119° E.	0.70960 ± 5	0.51298 ± 2
D-11 Amid. 119° E.	0.71384 ± 5	0.51305 ± 3

*Eimer and Amend SrCO_3 , $^{87}\text{Sr}/^{86}\text{Sr} = 0.70808$.

†Normalised to $^{146}\text{Nd}/^{144}\text{Nd} = 0.7219$.

preliminary study has now been extended in the light of the estimates of the $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of seawater as inferred from the studies of authigenic ferromanganese deposits. $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios have been measured in holocrystalline basalts from DSDP Leg 38 (North Atlantic Ocean) and dredged basalts from the south-east Indian Ridge. These results (Table 2), together with those for DSDP Leg 37 samples^{23,37}, are compared with the range of $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in unaltered basalts in Fig. 3. The samples plot to the right of this field to higher $^{87}\text{Sr}/^{86}\text{Sr}$ ratios up to and exceeding the seawater value, but are not significantly modified in Nd-isotope composition. The source of Sr with a $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.71384 ± 5 in D-11 Amid (Table 2; Fig. 3) is uncertain but is probably from local near-axis sediment. However, even this sample retains a $^{143}\text{Nd}/^{144}\text{Nd}$ ratio typical of mid-ocean ridge basalt. Ocean floor basalt is thus an important sink for radiogenic marine Sr, but not for Nd. The isotopic signature of altered ocean-floor basalt is sufficiently distinct to render it recognisable in some island arc tholeiites^{38,39}. $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of South Sandwich Island arc-tholeiites are shown in Fig. 3 and reveal the involvement of altered ocean-floor basalt in their generation.

Source of Nd, Pb and Sr in seawater

The isotopic compositions of Sr, Nd and Pb in oceanic ferromanganese deposits reflect both the sources of these elements and the differences of their residence times in the oceans. The $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of the samples (with the exception of F-55) are close to the measured value for seawater, and $^{143}\text{Nd}/^{144}\text{Nd}$ ratios are similarly relatively constant. The isotopic composition of Pb in the samples is highly variable and markedly provenance dependent. Thus, whereas the oceans are a substantial (10^{17} kg) and apparently well-mixed reservoir for Sr, the presence of Pb in seawater is such a transient phenomenon that the oceans cannot be considered as a reservoir at all for this element. Rather, the isotopic composition of Pb at some point in the oceans most strongly reflects the inputs in that vicinity from either the suspended or deposited particulates (and possibly authigenic materials), or from hydrothermal additions from the basaltic oceanic crust. The situation with respect to Nd is a little less clear; there is apparently more uniformity in $^{143}\text{Nd}/^{144}\text{Nd}$ ratios and the oceans may be a small ($\sim 4 \times 10^9$ kg) but relatively well-mixed reservoir for Nd.

As far as the source of Sr in the oceans is concerned the situation has already been adequately discussed^{7,8,40}, and it is clear that the dominant source is continental weathering, and that the flux from mid-ocean ridges plays a very subordinate role. The basaltic part of the oceanic crust removes radiogenic ^{87}Sr from seawater either by isotope exchange or by addition of seawater Sr to the basalt and recycles continental Sr into the mantle. A proportion of this Sr seems to be returned to the surface at island arcs^{38,39}. In this way the oceans behave as an important intermediate reservoir through which continental Sr is cycled into the mantle. As indicated above, the concept of the oceans as a reservoir for Nd may be valid, and the continents and oceanic basalts (mantle) are likely sources

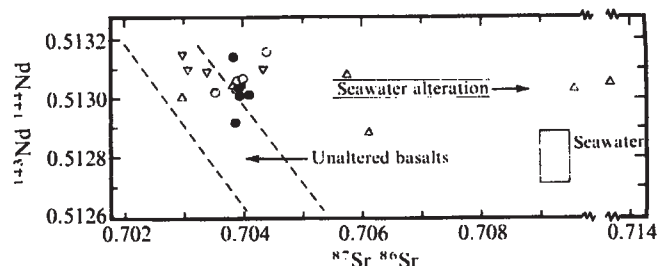


Fig. 3 Comparison of $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{87}\text{Sr}/^{86}\text{Sr}$ ratios in unaltered basalts from DSDP Leg 37 (Δ)²³, DSDP Leg 38 ($\circ$) and the south-east Indian Ridge ($\wedge$) (Table 2), with island-arc tholeiites from the South Sandwich Islands ($\bullet$)³⁰, and L-DGO unpublished data.

of supply. The $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of the mantle source can be estimated quite precisely, and is close to 0.5130 (Fig. 2). The $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of average continental material is more difficult to estimate but will be < 0.5120 (see also ref. 39); however, it is not known whether or not average continental Nd is being transported to the oceans: there is no *a priori* reason to suspect that it is. Even with these uncertainties it is clear that a substantial proportion of Nd in the oceans is of continental derivation, but until the $^{143}\text{Nd}/^{144}\text{Nd}$ ratio of 'continental' input into the oceans is known it is difficult to specify the amount precisely.

We thank W. S. Broecker for stimulating our interest in oceanic ferromanganese deposits and for numerous discussions. The manganese nodules were supplied by A. Greenslate. We thank G. Weisselman for her technical help. This study was supported by US NSF grant EAR-75-20891. L-DGO contribution number 2679.

Received 22 December 1977; accepted 22 March 1978.

- Bonatti, E. A. *Rev. Earth planet. Sci.* 3, 401 (1975).
- Bostrom, K. & Peterson, M. N. A. *Econ. Geol.* 61, 1258 (1966).
- Bostrom, K. & Peterson, M. N. A. *Mar. Geol.* 7, 427 (1969).
- Corliss, J. B. *J. geophys. Res.* 76, 8128 (1971).
- Bender, M. *et al. Earth planet. Sci. Lett.* 12, 425 (1971).
- Dasch, E. J., Dymond, J. R. & Heath, G. R. *Earth planet. Sci. Lett.* 13, 175 (1971).
- Spooner, E. T. C. *Earth planet. Sci. Lett.* 31, 167 (1976).
- Brass, G. W. & Turekian, K. K. *Earth planet. Sci. Lett.* 34, 165 (1977).
- Bender, M. L., Ku, T.-L. & Broecker, W. S. *Earth planet. Sci. Lett.* 8, 143 (1970).
- Turekian, K. K. *Geochim. cosmochim. Acta* 41, 1139 (1977).
- Bender, M. L. in *Ferromanganese Deposits on the Ocean Floor* (ed. Horn, D. R.) (IDOE Publ. NSF, 1972).
- Piper, D. Z. in *Ferromanganese Deposits on the Ocean Floor* (ed. Horn, D. R.) (IDOE Publ. NSF, 1972).
- Ehrlich, A. M. thesis, Massachusetts Institute of Technology (1968).
- Chow, T. J. & Patterson, C. C. *Geochim. cosmochim. Acta* 26, 263 (1962).
- Reynolds, P. H. & Dasch, E. J. *J. geophys. Res.* 76, 5124 (1971).
- Goldberg, E. D. *IAEA Publ.* 68, 121 (Vienna, 1963).
- Nozaki, Y. & Tsunogai, S. *Earth planet. Sci. Lett.* 20, 88 (1973).
- Schell, W. R. *Geochim. cosmochim. Acta* 41, 1019 (1977).
- Rama, K. M. & Goldberg, E. D. *Science* 134, 98 (1961).
- Nozaki, Y., Thomson, J. & Turekian, K. K. *Earth planet. Sci. Lett.* 32, 304 (1976).
- Craig, H., Krishnaswami, S. & Somayajulu, B. L. K. *Earth planet. Sci. Lett.* 17, 295 (1973).
- Garrels, R. M., Mackenzie, F. T. & Hunt, C. *Chemical Cycles and the Global Environment* (Kaufmann, Los Altos, 1975).
- O'Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* 34, 13 (1977).
- Arden, J. A. & Gale, N. H. *Analyt. Chem.* 46, 2 (1974).
- Veizer, J. & Compston, W. *Geochim. cosmochim. Acta* 38, 1468 (1974).
- Papanastassiou, D. A., Wasserburg, G. J. & Burnett, D. S. *Earth planet. Sci. Lett.* 8, 1 (1970).
- Hildreth, R. A. & Henderson, W. T. *Geochim. cosmochim. Acta* 35, 235 (1971).
- DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* 3, 249 (1976).
- DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* 3, 743 (1976).
- Carter, S. R., Evensen, N. M., Hamilton, P. J. & O'Nions, R. K. *Earth planet. Sci. Lett.* 37, 401 (1978).
- Church, S. E. & Tatsumoto, M. *Contr. Miner. Petrol.* 53, 253 (1975).
- Cumming, G. L. *Earth planet. Sci. Lett.* 31, 179 (1976).
- Stacey, J. S. & Kramers, J. D. *Earth planet. Sci. Lett.* 26, 207.
- Cumming, G. L. & Richards, J. R. *Earth planet. Sci. Lett.* 28, 155 (1975).
- Dasch, E. J., Hedge, C. E. & Dymond, J. *Earth planet. Sci. Lett.* 19, 177 (1973).
- Hart, S. R., Erlank, A. J. & Kable, E. J. D. *Contr. Miner. Petrol.* 44, 219 (1974).
- O'Nions, R. K. & Pankhurst, R. J. *Earth planet. Sci. Lett.* 31, 255 (1976).
- Hawkesworth, C. J., O'Nions, R. K., Pankhurst, R. J., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* 36, 253 (1977).
- DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* 4, 465 (1977).
- Brass, G. W. *Geochim. cosmochim. Acta* 40, 721 (1976).

Physical mechanisms for chemically sensitive semiconductor devices

P. Bergveld & N. F. DeRoij

Twente University of Technology, Department of Electrical Engineering, Enschede, The Netherlands

J. N. Zemel

University of Pennsylvania, Department of Electrical Engineering and Science, Moore School of Electrical Engineering, Philadelphia, Pennsylvania 19174

The thermodynamics of semiconductor based chemical sensors is examined. Depending on the structure, both ionic and neutral species can be detected. A reference electrode is required only if an ionic species is involved and there is no interaction by surface states at the semiconductor-sensor material interface. If any chemical process at the solution-sensor interface frees a diffusing species which interacts with the semiconductor interface, then it may be feasible to dispense with an external reference and use the internal Fermi level of the semiconductor as the reference. A parallel argument is presented for neutral species leading to the well known Henry's law for gas phase doping of semiconductors. In both neutral and ionic detection, rapid forward and reverse processes are essential for 'reversible' detection.

SEVERAL general papers have been published recently on the potential of chemically sensitive semiconductor devices (CSSDs)¹⁻³. The solid information on the operation of a limited class of devices is sufficient to establish their promise⁴⁻⁶, but there has been much discussion about the conceptual basis for these devices that is largely rooted in the different languages of electrochemistry and device physics^{3,7,8}. Ion-selective electrodes have been the principal solid state devices used for chemical sensing and much is known about these electrodes⁹⁻¹². It has been natural to use the device physics of these structures as the conceptual basis for the ion selective field effect transistor (ISFET) as well as other CSSDs. In general, solid-state chemical sensors respond to changes in the external chemical environment with a change in an electrical property or parameter such as conductivity, polarisation or refractive index. These changes occur in the chemically sensitive portion of the device and are detected by an electronic structure in intimate contact with the chemically sensitive region. The ISFET structure illustrates

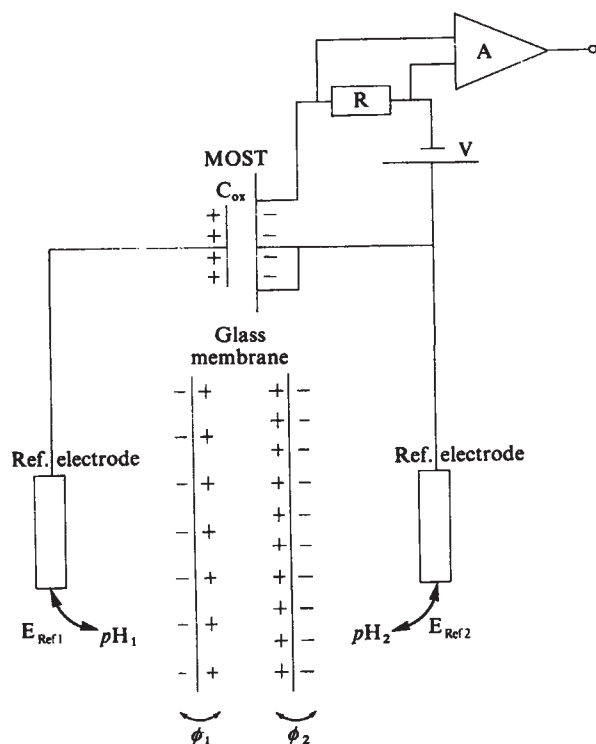


Fig. 1 Schematic diagram of a glass membrane, ion selective electrode with associated field effect transistor electrometer.

this configuration³⁻⁶. It is this composite character of the CSSD that makes it possible to envisage substantially new families of CSSD structures. In this article we examine the fundamental operation of the CSSD and emphasise the importance of thermodynamic driving forces for understanding the device operation. Some illustrative examples are presented which describe how the ISFET does operate.

The pH glass membrane electrode

To put matters in context, we quickly review the principles underlying the operation of the glass membrane electrode. This electrode is a simple one port system¹³ consisting of an ion permeable glass membrane separating two solutions. One solution has a constant ionic composition (usually the inner reference solution) and the other is a test solution; both solutions are contacted by standard reference electrodes (Fig. 1). Hydrogen ions cross both the liquid-glass interfaces and the associated hydrated gel layers separating the solutions. The reversibility of the exchange of protons in solution with those in the glass phase makes these electrodes reversible. An ionic double layer arises at the glass-solution interface which produces the junction potential. These potentials are due in part to diffusion of ionic charges within the glass (diffusion potential) and ion exchange at the surface (boundary potential). The resulting electromotive force (emf) of the cell can be measured between the two reference electrodes with an accuracy of 0.01 pH units if the measuring device has an internal resistance at least 10³ times greater than the internal resistance of the glass membrane. This resistance is determined by the resistivity and thickness of the relatively dry glass bulk region. Because the glass electrode is a one port system, there must be a continuous circuit through the entire system. The transport process through the bulk of the glass membrane is largely ionic in character.

At the reference electrode-solution interface there is a reversible charge exchange which provides electronic carriers for the external circuit. The ionic transport mechanism in pH selective glasses has been shown to be due to sodium ion motion¹⁴. There is essentially no perturbation of the glass composition because the mobile sodium ions are a small fraction of the total sodium concentration in the glass. Taking the

typical thickness of the glass electrode as 1 mm, it is simple to show that there are approximately 10²¹ sodium atoms cm⁻² in the glass. A current density of 10⁻¹⁰ A cm⁻² transports about 2 × 10¹² ions cm⁻² in 1 h, a negligible fraction of the total number of sodium atoms. However, this current must be several orders of magnitude smaller than the zero bias exchange currents in order to prevent electrode polarisation at the glass-liquid interface. The various electrochemical reactions at the different interfaces have to be reversible in character to assure proper operation of the devices. Most standard electrometers use a MOSFET preamplifier where the input is connected directly to the gate electrode as shown in Fig. 1. The junction potential differences, φ₁(pH₁) - φ₂(pH₂), between the two glass-solution interfaces can be readily measured, as it charges the gate capacitor of the MOSFET. The strength of the inversion layer is measured in an entirely conventional fashion from the conductance of the inversion layer.

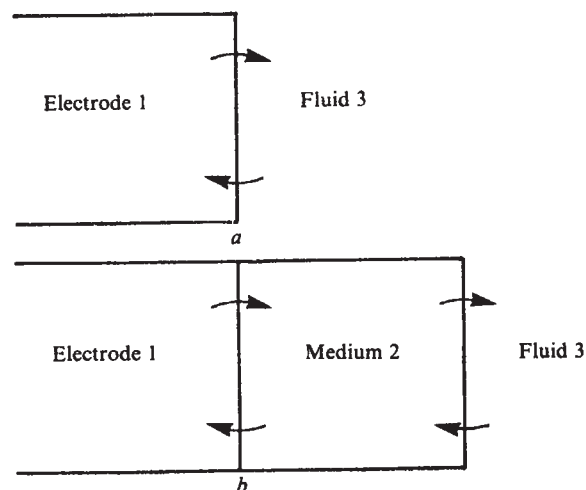
The generalised CSSD

In order to understand the ISFET and other CSSDs, it is best to generalise these devices as shown in Fig. 2a. The electrode in Fig. 2a may have an internal structure that permits us to determine the electrical properties of the electrode without reference to the external fluid. Chemical constituents in the fluid exchange with the electrode through surface adsorption and diffusion. In the language of thermodynamics, the driving forces are the differences in chemical potentials in the case of neutral transportable species and electrochemical potentials in the case of charged transportable species within the fluid and the electrode. At equilibrium, the general condition that the (electro)chemical potentials of the transportable species between the fluid and the electrode must be equal, must apply. At equilibrium, the Fermi-level within the CSSD electrode can be evaluated and, as a result, can be related to changes in the composition of the fluid. In other words, if $\tilde{\mu}_1$ is the electrochemical potential (ECP) of the exchangeable component in the CSSD electrode, and $\tilde{\mu}_3$ is the ECP of the corresponding component in the fluid, then, in equilibrium conditions

$$\Delta \text{ECP} = \Delta \tilde{\mu}_1 = \Delta \tilde{\mu}_3 \quad (1)$$

in which Δ indicates a change. The requirements of bulk charge

Fig. 2 Schematic diagram of chemically sensitive semiconductor devices. a, Simple structure where there is a direct interaction between the fluid and the device. b, Composite structure where medium 2 represents a chemically responsive layer that can interact with the device and the fluid both.



neutrality and extrinsic behaviour for an n -type semiconductor electrode yield

$$n = N_D - N_A \quad (2)$$

where n is the electron carrier concentration, N_D is the concentration of donor impurities and N_A is the concentration of acceptor impurities. Since the change in electrochemical potential of the semiconductor is

$$\Delta\tilde{\mu}_1 = kT \ln(n_f/n_i) \quad (3)$$

where n_f and n_i are the final and initial electron concentrations when the system is in its final and initial states, respectively, then we find that

$$\Delta\tilde{\mu}_1 = kT \ln(n_f/n_i) = kT \ln[(N_{Df} - N_{Af})/(N_{Di} - N_{Ai})] \quad (4)$$

where the f and i subscripts are as above. If we assume that $N_{Df} \gg N_{Af}$ and $N_{Di} \gg N_{Ai}$ then

$$\Delta\tilde{\mu}_1 = kT \ln(n_f/n_i) \approx kT \ln(N_{Df}/N_{Di}) \quad (5)$$

The fluid phase will have an activity of the donor species, a_D , which corresponds to

$$\Delta\tilde{\mu}_3 = kT \ln(a_{Df}/a_{Di}) \quad (6)$$

Combining equations (1), (5) and (6) yields

$$n_f/n_i = N_{Df}/N_{Di} = a_{Df}/a_{Di} \quad (7)$$

This is analogous to Henry's law for ideal solutions and is essentially identical to discussions of doping of semiconductors from the gas phase¹⁷.

There are several important implications of equation (7). First, there must be a transport process which allows the fluid-electrode system to move from one equilibrium state to another. Diffusion processes will be governed by the differences in the chemical potential of the transportable species and the boundary conditions existing at the fluid-electrode interface. In the case of charged transportable species, differences in the electric potential will arise across the interface. If we focus our attention on the interface, it becomes clear that there is no reason why this region has to be infinitesimally thin. If we can replace the interface with a film as indicated in Fig. 2b, the flexibility of the CSSD concept is increased substantially.

Instead of an adsorption process being the intermediate state through which equilibrium is achieved, a more complex process of adsorption and internal diffusion can take place in the film and the transportable species need not interact directly with the electrode. Instead, the transportable species can interact with the film and produce substantial changes in the electrical properties of the film alone. The electrode can be a device with a structure of more or less conventional design¹⁸.

Equation (7) will still apply to the interaction of a semi-conducting or insulating film with an external fluid. If the film-electrode structure forms a hetero- or Schottky barrier junction (Fig. 2b, interface 1-2), then the change in the Fermi level of the film ($\tilde{\mu}_2$) (or work function if this is a more appropriate concept) is reflected in a change in the barrier height across the junction 1-2. This has been adequately discussed by Zemel¹⁵. If an ion exchange process occurs at the film-fluid interface 2-3 which changes $\tilde{\mu}_2$, then the induced polarisation at 1-2 can be measured¹⁸. The simplest way of measuring the polarisation is with a capacitor consisting of the electrode-film structure and another (in some cases, reference) electrode. This change in polarisation may be reflected in changes in surface charge or admittance of the capacitor.

Because the chemical sensing function is separate from the electronic device structure, such devices will involve staged phenomena, that is, in the first stage the film and the fluid interact by mass diffusion resulting from chemical processes at the 2-3 interface and in the second stage, the diffused particles interact at the 1-2 interface to change the work function difference, ϕ_{12} . The change in ϕ_{12} has direct and obvious consequences since the barrier height between the film and the electrode is substantially ϕ_{12} .

As is well known in semiconductor surface electronics, surface trapping states introduce an additional polarisation into the MOS type system. This charge is in parallel with the semiconductor space charge region and, since charge neutrality must still be preserved, we can write

$$Q_{ss} + Q_{sc} + Q_{film} = 0 \quad (8)$$

where $Q_{ss} = qN_{ss}$ is the surface state charge, Q_{sc} is the semiconductor space charge and Q_{film} is charge in the film. The surface trapping state density arises from many factors. One process that has been widely considered is where an impurity atom migrates from the 2-3 interface to the 1-2 interface and becomes electrically active at that interface. Quantitatively we can write the number of negatively charged surface trapping levels as

$$N_{ss}^- = \int_E \frac{N_{ss}(E)dE}{1 + \exp\{(E - E_f)/kT\}} \quad (9)$$

where $N_{ss}(E)$ is the energy dependence of the surface state density. If we expand $N_{ss}(E)$ in powers of E , we obtain

$$N_{ss}(E) = N_{ss}^0 + \left. \frac{\partial N_{ss}}{\partial E} \right|_{E=E_{ss}} (E - E_{ss}) + \frac{1}{2} \left. \frac{\partial^2 N_{ss}}{\partial E^2} \right|_{E=E_{ss}} (E - E_{ss})^2 + \dots \quad (10)$$

and calling

$$(1/n!) \left. \frac{\partial^n N_{ss}}{\partial E^n} \right|_{E=E_{ss}} = b_n \quad (11)$$

we obtain

$$\begin{aligned} N_{ss} &= \sum_E \int \frac{b_n (E - E_{ss})^n dE}{1 + \exp\{(E - E_f)/kT\}} \\ &= \sum b_n (kT)^{n+1} F_n\{(E_f - E_{ss})/kT\} \end{aligned} \quad (12)$$

where F_n is the Fermi function of n th order. Using the usual definition of the surface potential, we can write¹⁹

$$Q_{ss} = -(Q_{ox} + 2qn_i L_D F(u_s, u_b)) = q \sum b_n (kT)^{n+1} F_n(u_{ss}) \quad (13)$$

where n_i is the intrinsic carrier concentration of the semiconductor, L_D is the intrinsic Debye length, $F(u_s, u_b)$ is the charge function¹⁸, u_b is the bulk potential in kT and the u_i are corresponding surface state (ss) and surface (s) potentials. To determine the actual surface charge, we would have to extract the value of u_s that would satisfy equation (13). This value of u_s would be a function of both u_b , u_{ss} and N_{ss} . In principle, both u_{ss} and N_{ss} will be dependent on the concentration of

atoms or molecules freed by the chemical processes to be discussed below. However, this concentration of released atoms or molecules at the semiconductor surface will depend on diffusion in the film and the chemical kinetics associated with the 'annealing' or formation of surface states. Thus a combination of both chemical and physical measurements are needed to study the mechanisms proposed here.

The equilibrium electrochemical processes occurring in the CSSD's described by Fig. 2, can be illustrated with the following examples. In the case of Fig. 2a, the following two possibilities can arise when the electrode M is in contact with the fluid f: (1) if the electrode (M) reacts as



this will lead to the relationship:

$$\mu_M = \mu_M^0 + zq\psi_f + kT \ln a_{M_f^{z+}} + z\mu_{eM} - zq\psi_M \quad (15)$$

for $\psi_M - \psi_f$ as a function of the activities of the chemical species.

(2) If the electrode is inert, but the fluid contains a redox couple²⁰



and the equilibrium



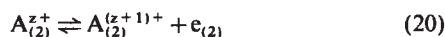
is established, the following relations can be derived:

$$\mu_{A_f^{z+}} + kT \ln a_{A_f^{z+}} = \mu_{A_f^{(z+1)+}} + kT \ln a_{A_f^{(z+1)+}} + \mu_{e_f} \quad (18)$$

and

$$\mu_{e_f} - q\psi_f = \mu_{e_M} - q\psi_M \quad (19)$$

Applying these two cases for the system shown in Fig. 2b, it is clear that phase (2) should now contain a redox couple

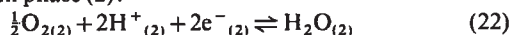


while at the interfaces 1-2 and 2-3 respectively the following reversible reaction should be occurring:



Note that A can represent either a charged or uncharged species. As an example, the following equation may illustrate the theory.

Redox couple in phase (2):



equilibrium at 2-3: $O_{2(2)} \rightleftharpoons O_{2(f)}$

(23)

equilibrium at 1-2: $e_{(2)} \rightleftharpoons e_M$

At a constant H^+ and H_2O activity, $\tilde{\mu}_{e(2)}$ depends only on the O_2 activity in the fluid and thus also $\tilde{\mu}_{eM}$.

Ion-sensitive field effect transistors

The liquid gate transistor operates in the same fashion as a MOSFET with a solution generally replacing the metal gate. Considering the oxide as the film as discussed above, the requirement for reversibility and induced polarisation is still stringent. The valinomycin-PVC membrane used by Moss, Janata and Johnson⁶ for K^+ , the Si_3N_4 layer by Esashi and Matsuo²¹ and the aluminised SiO_2 layer used by de Rooij and Bergveld²² for H^+ are both polarisable and reversible. Buck and Hackleman¹⁸ used an AgBr layer on a MOS capacitor to monitor the behaviour of the polarisable electrode. In all these the electrode becomes polarised and the resulting displacement currents will charge the ISFET (Fig. 3).

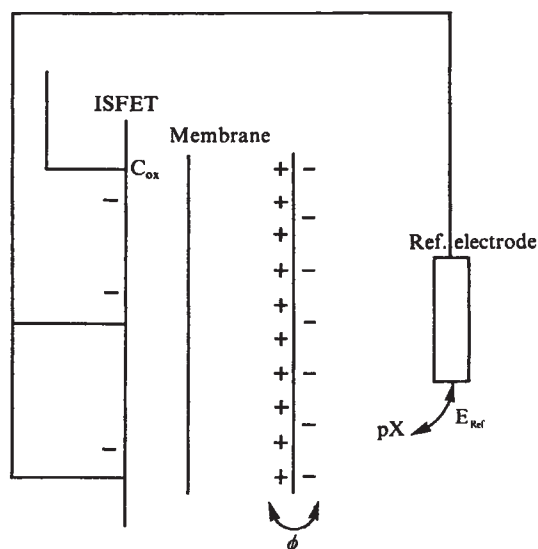
The OSFET capacitor consists of the film (Fig. 2b) and any insulation on the electrode, such as the gate oxide and nitride. However, the device will ideally respond only to the charging of the ion-exchange layer, arising from polarisation of the interface 2-3. If the silicon dioxide itself has ion sensitivity and selectivity properties there would be no need for the ion exchange membrane and the device will operate as an ISFET. In practice, pure silicon dioxide does not have ion-selective properties but requires a top layer of Si_3N_4 to produce pH selectivity^{5,21}. The operation of Si_3N_4 ISFETs can be readily explained with the above argument. No ionic transport is necessary through the bulk of the Si_3N_4 surface layer: only a specific polarisation to provide the proton sensitivity and selectivity and a reference electrode for charge transport through the external circuit. It is simple to show from Fig. 4, that the charge on the semiconductor device, ΔQ_{sc} , is but a fraction of the charge induced in the ion exchange layer, ΔQ_{DL} . The result is

$$\Delta Q_{sc} = [C_{ox}/(C_{ox} + C_{DL})]\Delta Q_{DL} \quad (24)$$

where $C_{sc} \gg C_{ox}$. We assume a value of $10 \mu F cm^{-2}$ for C_{DL} and $70 nF cm^{-2}$ for C_{ox} (500 Å silicon dioxide). This leads to $\Delta Q_{sc} = 7 \times 10^{-3} \times \Delta Q_{DL}$. Therefore, only 0.7% of the negative charge will be located in the silicon. Since the potential drop across C_{DL} is of the order of 58 mV per decade of adsorbed or ion-exchanged H^+ ions, the charge variation at this double layer will be about $5.8 \times 10^{-7} C cm^{-2}$ per decade. This means that $\Delta Q_{sc} \approx 3.5 \times 10^{-9} C cm^{-2}$ per decade, which is quite measurable.

The ISFET can also be exactly described thermodynamically in a manner similar to the methods used above. Thus if we assume that the silicon is contacted by silver and that the

Fig. 3 Schematic diagram of an ion-selective field effect transistor.



silicon oxide has a selective and sensitive top layer which can be treated as a single layer, the system simplifies to:

Ag	AgCl	Chloride solution with constant chloride activity and variable proton activity	Oxide	Silicon	Ag
1	2	3	4	5	1

Interface 4-5 is assumed to be ideally polarisable, whereas interfaces 1-2, 2-3 and 3-4 are reversible and can be described as follows



From these equilibria it follows that:

$$\tilde{\mu}_{e_{Ag}} + \mu_{AgCl} = \mu_{Ag} + \tilde{\mu}_{Cl^-} \quad (28)$$

$$\tilde{\mu}_{H^+_{sol}} = \tilde{\mu}_{H^+_{ox}} \quad (29)$$

$$\tilde{\mu}_{e_{Ag}} = \tilde{\mu}_{e_{Si}} \quad (30)$$

where $\tilde{\mu}$ is an electrochemical potential and μ a chemical potential.

From equations (28) and (30) it follows that

$$\tilde{\mu}_{e_{Si}} + \mu_{AgCl} = \mu_{Ag} + \tilde{\mu}_{Cl^-} \quad (31)$$

$$\text{where } \tilde{\mu}_{Cl^-} = \mu_{Cl^-}^0 + kT \ln a_{Cl^-} - q\phi_{sol} \quad (32)$$

$$\text{and } \tilde{\mu}_{e_{Si}} = \mu_{e_{Si}}^0 - q\phi_{Si} \quad (33)$$

The potentials ϕ are the bulk values.

It now follows that

$$\begin{aligned} \phi_{sol} - \phi_{Si} &= (\mu_{Ag} - \mu_{AgCl} - \mu_{Si}^0 + \mu_{Cl^-}^0)/q + (kT/q) \ln a_{Cl^-} \\ &= \text{constant} = C_1 \end{aligned} \quad (34)$$

because the chloride activity is assumed to be constant. In the same way it follows from equation (26) that

$$\phi_{ox} - \phi_{sol} = C_2 + (kT/q) \ln \{(a_{H^+})_3/(a_{H^+})_4\} \quad (35)$$

where the subscripts refer the solution (3) oxide (4) regions. If H^+ is assumed to be the potential determining ion then $(a_{H^+})_4$ is constant by definition²³. Then from equation (34) and (35), it follows finally that

$$\phi_{ox} - \phi_{Si} = C_1 + C_2 + (kT/q) \ln(a_{H^+}) \quad (36)$$

which is the voltage across the oxide capacitance C_{ox} . It may thus be concluded that the corresponding charge which loads the oxide capacitance is fully thermodynamically determined.

Discussion and conclusions

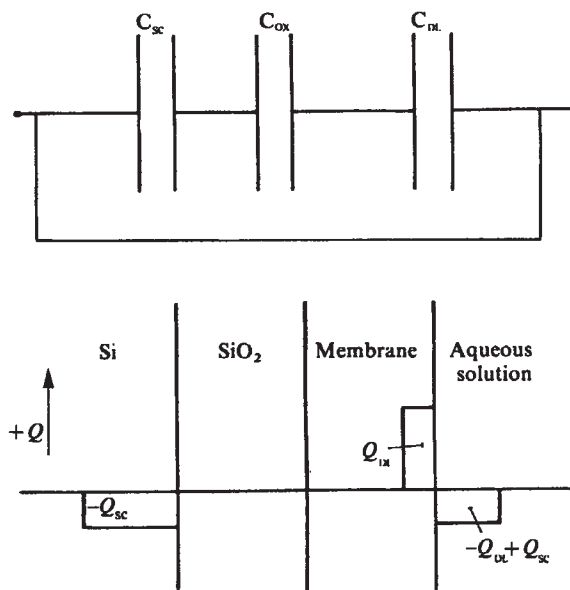
There are two fundamental conclusions to be drawn from the above results. First, the chemical driving forces resulting from reversible reactions at the electrode-fluid interface (Fig. 2a) provide the conceptual framework for understanding the CSSD. Second, the signal arises from physical phenomena which respond to normal chemical processes. In general the electrode will have an internal structure which generate signals proportional to changes in the electrochemical potential of the sensitive medium. We have shown that when an electrically active component of the fluid enters a semiconductor or insulator, we obtain the equivalent of Henry's law for the ratio of the carrier concentration in the sensitive film and the activity of the component in the fluid. This result is the thermodynamic origin of the device equations proposed by Zemel earlier¹⁶.

If the internal electronic device structure of the electrode is separate from the chemically active region, then there must be a mechanism providing transport from the sensing region to the electronic device structure. The two mechanisms proposed are displacement currents in what is essentially a dielectric constant measurement, or the modification of surface state density by species generated at the sensor-fluid interface and transported by diffusion to the sensor-electronic device interface. This latter process is analogous to the annealing of surface states by forming gas. The former process is perhaps more general and allows the experimenter to select a broad range of polarisable materials as part of a sensor element (SE). Among these layers we would include ionic crystals, impregnated polymers and sputtered ion sensitive glasses. Much will depend on the extent of the attachment of the SE to the silicon oxide, their thickness and electrical properties. If the specific layer has an insulating bulk and is only ion sensitive due to specific adsorption (for instance Si_3N_4), then the sensitivity of the device is a direct function of the membrane thickness.

Therefore these materials should be very thin, for instance, 100 Å. Polymer layers like PVC and silicone rubber, where only the top layer is impregnated with ion-sensitive materials are usually of the order of tens to hundreds of μm in thickness. This could reduce their sensitivity due to voltage division. If the SE on the silicon oxide is conductive, then thicker layers can be used.

For further development of ion selective semiconductor devices membrane materials will be needed that will adhere tenaciously to the semiconductor surface. Ion-exchange materials are also needed that can be included in the membrane

Fig. 4 Capacitance and charge schematic for an ISFET.



and still retain their desirable properties and be compatible with planar processing. Isolation of leads and preprocessing electronics from the hostile environment represented by saline solutions, acid baths and so on is of fundamental technological importance. The development of two sided structures for ISD's is now possible as a result of the research of Anthony and Cline on gradient zone melting processes²⁴. The compatibility of membrane formation with planar processing still has to be addressed. Until this problem is fully solved, the multi-element ion selective device will continue to be a laboratory curiosity rather than a major analytic tool for the chemical sciences.

We conclude that there is a need for substantial research on the microscopic interactions at the fluid-solid interface. In particular kinetic studies of these interactions are needed to advance the subject.

J.N.Z. thanks Dr T. C. Chen and B. Keamati for useful discussions and Dr I. Lauks for a careful reading of the manuscript, and the Technicon Instrument Company, the Pennsylvania Science and Engineering Foundation and the Naval Research Laboratory for their support. N.F.de R. and P.B. thank the Netherlands Organisation for the Advancement of Pure Research (ZWO) for their support, and Professors J. Th. G. Overbeek and M. J. Sparnaay for useful discussions.

Received 30 January; accepted 20 March 1978.

1. Zemel, J. N. *Analyt. Chem.* **47**, 255A (1975).
2. Zemel, J. N. *Res. Dev.* **28**, 38 (1977).
3. Kelly, R. G. *Electrochim. Acta* **22**, 1 (1976).
4. Bergveld, P. *IEEE Trans. Biomed. Engng BME-19*, 342 (1972).
5. Matuso, T. & Wise, K. D. *IEEE Trans. Biomed. Engng. BME-21*, 485 (1974).
6. Moss, S. D., Janata, J. & Johnson, C. C. *Analyt. Chem.* **47**, 2238 (1975).
7. Bergveld, P. *Biomed. Engng* **11**, 359 (1976).
8. Janata, J. & Moss S. D. *Biomed. Engng* **11**, 241 (1976).
9. Rechnitz, G. A. *Science* **190**, 234 (1975).
10. Bates, R. G. *Determination of pH: Theory and Practice* (Wiley, New York, 1973).
11. Eisenman, G. (ed.), *Glass Electrodes for Hydrogen and Other Cations* (Dekker, New York, 1967).
12. Durrst, R. A. (ed.) *Ion Selective Electrodes*, NBS spec. Publ. 314 (1969).
13. Desoer, C. A. & Kuh, E. S. *Basic Circuit Theory* (McGraw-Hill, New York, 1969).
14. Schwabe, K. & Dahms, H. Z. *Electrochem.* **65**, 518 (1961).
15. Zemel, J. N. *Intern. Electron Devices Meeting*, Washington, D.C., 635 (1975).
16. Keramati, B. & Zemel, J. N. *AEMB*, Los Angeles (1977).
17. Grove, A. S. *Physics and Technology of Semiconductors Devices* (Wiley, New York, 1967).
18. Buck, R. P. & Hackleman, D. E. *Analyt. Chem.* (in the press).
19. Kingston, R. H. & Neustadter, S. F. *J. appl. Phys.* **26**, 718 (1955).
20. Gerischer, H. *Adv. Electrochem. Electr. Engng* **1**, 139 (1960).
21. Esashi, M. & Matsuo, T. *Suppl. J. Japan Soc. appl. Phys.* **44** (1975).
22. Bergveld, P. & de Rooij, N. F. *Theory, Design and Biomedical Application of Solid State Chemical Sensors* (CRC Press, Cleveland, 1977).
23. Ahmed, S. M. *J. Phys. Chem.* **73**, 3546 (1969).
24. Cline, H. E. & Anthony, T. R. *J. applied Phys.* **47**, 2325 (1976).

The α -helix dipole and the properties of proteins

W. G. J. Hol, P. T. van Duijnen & H. J. C. Berendsen

Department of Chemistry, University of Groningen, Nijenborgh 16, 9747 AG, Groningen, The Netherlands

Phosphate moieties bind frequently at N-termini of helices in proteins. It is shown that this corresponds with an optimal interaction of the helix dipole and the charged phosphate. This favourable arrangement may have been discovered several times during evolution. In some enzymes, the helix dipole might be used in catalysis.

THE number of proteins of which the three-dimensional structure is known is increasing rapidly. The comparison of these known conformations, their amino acid sequences and ligand (substrate or coenzyme) binding sites has become an important new aspect of molecular biology. Focusing on the ligand binding properties, it seems that the negatively charged phosphate moieties of coenzymes and substrates are frequently bound at the N terminus of a helix (Table 1). This phosphate position allows a favourable interaction of its charge with the considerable dipole moment of the helix backbone¹. This suggests that several proteins utilise a 'phosphate binding helix' or 'nucleotide binding helix' as a tool for binding phosphate containing compounds. In addition, some enzymes may use an 'active site helix' to enhance reaction rates.

The helix dipole and its field

The α -helix dipole finds its origin in the dipoles of the single peptide units (Fig. 1). The formal charges given in Fig. 1 lead to a dipole moment of 3.5 D (1.2×10^{-29} Cm). This value is supported by experimental evidence of various kinds, as has been recently reviewed by Wada¹. In an α helix the peptide dipole moments are aligned nearly parallel to the helix axis, the axial component being 97% of the dipole moment. Theoretical considerations indicate an increase in the peptide dipole moment by polarisation due to hydrogen bonds in the α helix, yielding values of up to 5 D (ref. 1).

In the following we have ignored this polarisation, with the possible consequence of an underestimation of the electric fields.

The field of a helix can be approximated by the field of a continuous line dipole with a dipole density of 3.5 D per 1.5 Å, or 0.8×10^{-19} C. The field of such a line dipole is equal to the field of a positive charge at the amino end and a negative charge at the carboxyl end, each of magnitude 0.8×10^{-19} C (one-half elementary charge). Thus the helix, whose charges are actually distributed among the backbone atoms, provides the effect of an isolated half unit charge, with the compensating charge separated from it by the full length of the helix. This is in contrast to actual charges in proteins which are in practically all cases either solvated or compensated at close proximity in a salt bridge.

In Fig. 2a the simplified potential of the line dipole is plotted as a broken line; the drawn curve gives the potential on the axis of a helix consisting of peptide dipoles at their actual positions in the α helix. The line dipole seems to be a good approximation. Figure 2b shows the influence of helix length on the potential at a distance of 5 Å from the terminal residue. In these plots a value of 2 for the dielectric constant has been chosen. An effect that has not been taken into account is the reaction field of the highly polarisable solvent, which tends to decrease the field near the molecular surface². However, many substrate or coenzyme binding sites are clefts or depressions at the surface, where the influence of the reaction field is not fully developed. The neglect of the reaction field may therefore be considered to be a reasonable approximation.

Summarising we may state that (1), an α helix has a considerable electric field. Its dipole runs from the C-terminal to the N-terminal, that is, the positive end is near to the amino terminus. (2) The strength of the field increases up to a helix length of about 10 Å (seven residues or two turns), whereafter further elongation has only a marginal effect (Fig. 2b).

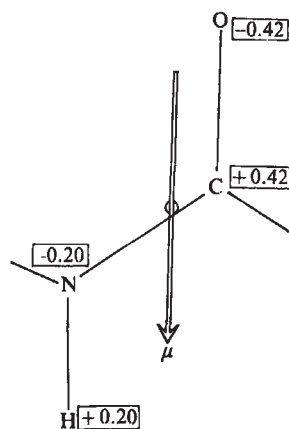


Fig. 1 Geometry and dipole moment of the peptide unit. Numbers in boxes give the approximate fractional atomic charges (in units of the elementary charge) according to ref. 2, with the charge of C changed into +0.42 in order to preserve charge neutrality. The dipole moment has a value of $0.72 \text{ eÅ} = 3.46 \text{ D} = 1.155 \times 10^{-30} \text{ Cm}$.

Possible effects of the α -helix dipole

The following possible effects will be considered. First, the helix field assists in binding of charged substrates or coenzymes. Negatively charged groups when bound at the amino terminus of the helix would interact favourably with the helix field. Second, the helix possibly causes long-range attraction of charged substrates. This possibility is particularly intriguing for those proteins which contain several helices running parallel to each other. And third, the helix field may contribute to the acceleration of reactions catalysed by enzymes. For instance, the transfer of protons away from side chains at the N-terminus of a helix would be facilitated.

Binding of charged groups at helix termini

The phosphate moieties listed in Table 1 are bound, typically, at a distance of 3–5 Å from the amino terminus of a helix. As shown in Fig. 2a, the potential at 5 Å from the N terminus of an α helix of 10 Å length is about 0.5 V. This means that binding of a single negatively charged group at that position involves an attractive energy contribution of 0.5 eV or about 12 kcal mol^{-1} . The free energy of desolvation per water molecule at a site 5 Å from the helix terminus, where the electric field is 10^9 V m^{-1} , is only about 200 cal mol^{-1} (estimated assuming a Langevin-type dipolar orientation of water dipoles in the helix field). Even if several water molecules have to be displaced from the end of the helix, the net effect is clearly attractive. This is particularly true when multiply-charged substances are bound.

Table 1 lists several proteins which bind negatively charged phosphate moieties of various compounds near the N terminus of a helix. The structures and amino acid sequences of these proteins vary widely, supporting our proposal for a role of the dipole moment of the helix backbone in these cases. As further evidence may be considered the fact that in three of these proteins no positively charged side chains are interacting with the phosphate groups. This is the case in flavodoxin (from *Clostridium* MP as well as from *Desulfovibrio vulgaris*⁴), lobster and bacterial glyceraldehyde phosphate dehydrogenase^{5,6} and yeast alcohol dehydrogenase⁷. Furthermore it is noteworthy that the positive charge on the nicotinamide moiety of NAD when bound to the dehydrogenases is situated quite far from the N-termini of the various helices occurring in these proteins (see for instance Figs 2 and 3 in ref. 8).

Long-range attraction of charged molecules

The possible involvement of helices in 'pulling' charged molecules to the protein molecule from the surrounding solvent is suggested by the structures of the coenzyme binding proteins⁸ lactate dehydrogenase, malate dehydrogenase, liver alcohol dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase and, in particular, flavodoxin. In all these instances, several helices run quite parallel to each other while the charged phosphates of the nucleotide are bound at the N terminus of one of those helices. Similarly, the structure determination of triose phosphate isomerase⁹ revealed several helices running approximately parallel to each other with their N termini orientated in such a manner with respect to the active site that the attraction of its negatively charged substrate could be facilitated.

The effects of such long-range cooperative helical fields are twofold. First, charged molecules are attracted to the binding site and, second, polar substrates or coenzymes are properly orientated by the field while they move towards the binding site. Although difficult to estimate quantitatively, both effects will increase the effective cross-section of the diffusion-controlled binding process and hence the bimolecular rate constant k_{on} . For physiological substrates the second-order rate constant generally is in the range 10^7 to $10^9 \text{ M}^{-1} \text{ s}^{-1}$ which is approximately the value expected for diffusion controlled processes¹⁰. Hence, the helix field will increase reaction rates when this diffusion process is rate limiting, that is, when the substrate concentration is lower than K_M . Interestingly, this is the concentration frequently observed in physiological conditions⁴.

Enhancement of reaction rates

Four enzymes in which the active site is located at the N terminus of an α helix are listed in Table 2. The mechanism of the two proteases mentioned in the table, papain and subtilisin, have several features in common^{12–15}. A group XH (SH in papain, OH in subtilisin) performs a nucleophilic attack on the substrate with a prior or simultaneous transfer of the hydrogen from XH to a group Y (an imidazole ring in both enzymes). The resulting covalent tetrahedral intermediate is negatively charged. The reaction proceeds then through further steps, including an acyl-enzyme, which need not be discussed here. The group XH is located at the amino terminal of a long central helix in these two, structurally completely unrelated, proteases. Two possible effects of the helix dipole on the catalysis may be

Table 1 Proteins binding negatively charged compounds at N termini of helices

Protein	'Phosphate binding helix'	Compound bound	Refs
Lactate dehydrogenase	34–36	NAD	31
Alcohol dehydrogenase	205–216	NAD	7
Glyceraldehyde phosphate dehydrogenase	13–25	NAD	5, 6
Flavodoxin	17–27	FMN	32, 33
Dihydrofolate reductase (<i>L. casei</i>)	99–107 & 42–49	NADP	*
Glyceraldehyde phosphate dehydrogenase	149–165	Glyceraldehyde-3-phosphate & inorganic phosphate	5, 6, 34
Triosephosphate isomerase	232–238	Glyceraldehyde-3-phosphate	9, 35
Phenylalanine t-RNA synthetase	helix E	Tyrosinyl adenylate	36†
p-Hydroxybenzoate hydroxylase	13–28	FAD	‡

*D. A. Matthews *et al.*: submitted to *J. biol. Chem.*

†Also, Monteilhet, C. & Blow, D. M., in preparation. In this case helix-phosphate distance appears to be about 10 Å.

‡R. K. Wierenga, K. H. Kalk, R. de Jong, W. H. & J. Drenth, unpublished.

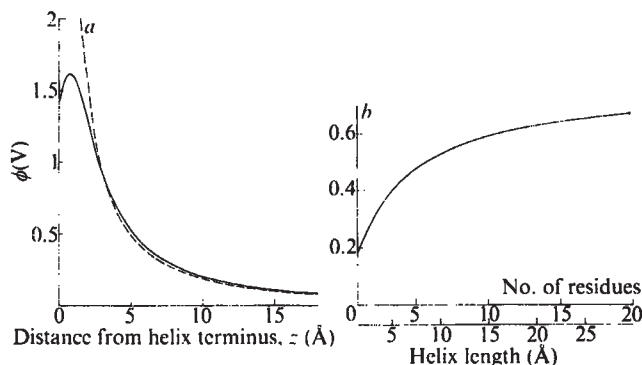


Fig. 2 *a*, Potential in volts due to a two-turn α -helix. Potentials are given on the helix axis as function of the distance z from the terminal dipole. Two models are used. Drawn curve: arrangement of 7 dipoles of 1.2×10^{-20} Cm on a helix at 1.9 Å from the axis with 1.5 Å translation and 100° rotation per dipole. Dashed curve: line dipole with homogeneous density of 8×10^{-20} C and length of 10.5 Å, corresponding to seven residues of an α helix. The latter potential is equivalent to that of a positive and a negative charge of 8×10^{-20} C at the ends of the line. Dielectric constant is 2 in both cases. *b*, Potential on the helix axis at 5 Å from the terminal dipole, for a helical arrangement of dipoles as a function of helix length.

considered: (1) facilitation of the transfer of the proton from XH to Y; (2) stabilisation of the tetrahedral intermediate. For a quantitative evaluation of the effect of the helix dipole on the proton transfer, the electrical potential along the line X–Y has been calculated for papain and subtilisin (Fig. 3). Quantum mechanical calculations (P.T. van D. *et al.*, in preparation) show that the change in dipole moment going from $XH \dots Y$ to $X^- \dots H^+Y$ is about 10D. Hence, given the field as depicted in Fig. 3, the latter arrangement is stabilised to an extent of about 10 kcal mol^{-1} . The helix clearly facilitates the transfer of the proton.

The effect of the helix dipole on the stabilisation of the transition state is hard to estimate. It is interesting, however, that in both proteases, the negatively charged carbonyl oxygen of the tetrahedral intermediate makes a hydrogen bond with the peptide NH of the first residue of the active site helix^{13,14}. The intermediate is consequently located very close to the N terminus of this helix, allowing a favourable interaction with the electrical field of the helix. In order to arrive at a reasonable estimate of the stabilising effect of the helix dipole, the charge distribution of the various species along the reaction path would have to be known in considerable detail. This is beyond the scope of this article.

A third enzyme from Table 2, glyceraldehyde phosphate dehydrogenase, shares several of the features discussed above for papain and subtilisin¹⁶. It contains, at the N terminus of a helix (Table 2), the unusually reactive sulphhydryl group of Cys 149 which is within hydrogen bonding distance of His 176. The orientation of the helix is such that proton transfer would be facilitated. This arrangement is very similar to that in papain and glyceraldehyde-3-phosphate dehydrogenase does indeed show esterase activity¹⁷. The facilitated proton transfer might also increase the rate of formation of a thiohemiacetal intermediate, proposed as a first step in the reaction with glyceraldehyde-3-phosphate¹⁷.

The fourth enzyme included in Table 2, rhodanese, is a sulphur transferase. In the first step of the reaction catalysed by this enzyme, the sulphhydryl group of Cys 247 performs a nucleophilic attack on the outer sulphur of its substrate thiosulphate^{18–20}. This essential cysteine is located at the N terminus of its active site helix. It seems likely that the

field of this helix is a factor involved in decreasing the pK of its sulphhydryl group²¹, and consequently in rendering this group more nucleophilic.

Discussion

In view of the evidence presented above, the functional role of helices in binding charged molecules seems to be well established. It is also likely that in some enzymes the field of the helix plays a part during catalysis.

Interestingly, chymotrypsin and related serine proteases, the active site of which resembles that of subtilisin very closely¹², do not contain any helix near their catalytic centres^{22–26}. Calculations show that, nevertheless, the electrical field due to the complete backbone as well as due to all atoms in chymotrypsin facilitates the transfer of the proton from its essential hydroxyl group to the nearby imidazole ring.

Table 2 Enzymes with the active site located at the N-terminus of a helix

Enzyme	'Active site helix'	Essential nucleophile	Refs
Papain	25–43	S γ of Cys 25	13, 14, 37
Subtilisin	221–237	O γ of Ser 221	12, 38, 39
Glyceraldehyde phosphate dehydrogenase	149–165	S γ of Cys 149	5, 6
Rhodanese	249–264	S γ of Cys 247	17, 18

The proposed role of the helix backbone is of importance for discussions concerning the evolutionary relationship of nucleotide binding proteins. The structural similarities of the 'nucleotide binding domain' in this group of proteins with quite unrelated sequences, have been considered possible results of a process of divergent evolution from a common ancestor^{8,27–30}. However, the frequent occurrence in

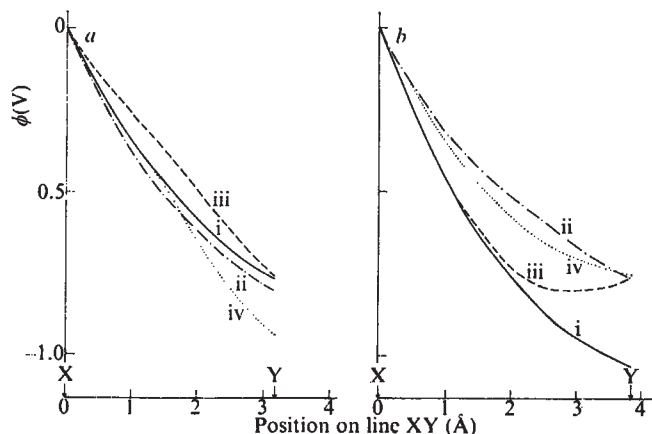


Fig. 3 Electrical potential on the line XY (see text) for two proteases. *a*, Papain: X = S γ of Cys 25, Y = N δ^1 of His 159. i, Backbone active site helix (24–43). ii, Backbone Cys 25 + all other atoms active site helix. iii, Backbone all residues. iv, Backbone Cys 25, His 159, Asn 175 + all other residues. *b*, Subtilisin: X = O γ of Ser 221, Y = N ϵ^2 of His 64. i, Backbone active site helix (220–238). ii, Backbone Ser 221 + all other atoms active site helix. iii, Backbone all residues. iv, Backbone Ser 221, His 64, Asp 32 + all residues. The coordinates of the O γ atom were taken from Matthews *et al.*⁴⁰. The potential at X was taken as (arbitrary) zero point. Partial atomic charges for calculations with complete residues were taken from Poland and Scheraga⁹. In order to simulate the effect of solvation, the charges of the exposed Asp, Glu, Arg and Lys side chains were reduced to half their normal values. For calculations using backbone only, the charges of Fig. 1 were used. The dielectric constant was taken $\epsilon = 1 + (R - R_A)/R$, R being the distance between a point on XY and a contributing atom, $R_A = 1.5$ Å. For $R - R_A < 0$, $\epsilon = 1$.

these proteins of the phosphates of the nucleotides near N termini of helices is remarkably consistent with a functional role of these helices. Consequently, the structural similarity could well be a result of convergent evolution towards an arrangement in which an optimal interaction exists between the charge of the coenzyme and the dipole of the 'nucleotide binding helix'.

In conclusion, we suggest that the electrical field generated by the backbone of a protein molecule, and by the helices in particular, is a significant factor which must be included in the discussion of the properties of proteins.

We acknowledge Professor J. Kraut's initial suggestion to look into the possible role of the dipole moment of helices. We thank Professor J. Drenth for stimulating discussions. The protein coordinates were kindly provided by Dr Richard J. Feldman. This study was carried out, in part, under the auspices of the Netherlands Foundation for Chemical Research (S.O.N.), with financial aid from the Netherlands Organisation for the Advancement of Pure Research (Z.W.O.).

Received 6 February; accepted 10 March 1978.

1. Wada, A. *Adv. Biophys.* **9**, 1-63 (1976).
2. Poland, D. & Scheraga H. A. *Biochemistry* **6**, 3791-3800 (1967).
3. Berendsen, H. J. C. *Report on CECAM Workshop Models in Protein Dynamics* 299-304 (CECAM, Orsay, France, 1976).
4. Mayhew, S. G. & Ludwig, M. L. *Enzymes* **12**, 57-118 (1975).
5. Moras, D. *et al.* *J. biol. Chem.* **250**, 9137-9162 (1975).
6. Biesecker, G., Ieuan Harris, J., Thierry, J. C., Walker, J. E. & Wonacott, A. J. *Nature* **266**, 328-333 (1977).
7. Brändén, C. I., Jörnvall, H., Eklund, H. & Furugren, B. *Enzymes* **11**, 103-190 (1975).
8. Rossmann, M. G., Liljas, A., Brändén, C. I. & Banaszak, L. J. *Enzymes* **11**, 61-102 (1975).
9. Banner, O. W. *et al.* *Nature* **255**, 609-614 (1975).

10. Hammes, G. G. & Schimmel, P. R. *Enzymes* **2**, 67-114 (1970).
11. Kraut, J. A. *Rev. Biochem.* **46**, 331-358 (1977).
12. Robertus, J. D., Kraut, J., Alden, R. A. & Birktoft, J. J. *Biochemistry* **11**, 4293-4303 (1972).
13. Glazer, A. N. & Smith, E. L. *Enzymes* **3**, 501-546 (1971).
14. Drenth, J., Kalk, K. H. & Swen, H. M. *Biochemistry* **15**, 3731-3738 (1976).
15. Drenth, J., Swen, H. M., Hoogenstraaten, W. & Sluyterman, L. A. A. E. *Kon. Nederl. Akad. Wetensch.* **78**, 104-110 (1974).
16. Garavito, R. M., Rossmann, M. G., Argos, P. & Eventoff, L. W. *Biochemistry* **16**, 5065-5071 (1977).
17. Ieuan Harris, J. & Waters, M. *Enzymes* **13**, 1-49 (1976).
18. Westley, J. *Adv. Enzym.* **39**, 327-368 (1973).
19. Ploegman, J. H. *et al.* *Nature* **273**, 124-129 (1978).
20. Ploegman, J. H. *et al.* *Nature* **273**, 124-129 (1978).
21. Schlesinger, P. & Westley, J. *J. biol. Chem.* **249**, 780-788 (1974).
22. Birktoft, J. J. & Blow, D. M. *J. molec. Biol.* **68**, 187-240 (1972).
23. Shotton, D. M. & Watson, H. C. *Nature* **225**, 811-816 (1970).
24. Stroud, R. M., Kay, L. M. & Dickerson, R. E. *J. molec. Biol.* **83**, 185-208 (1974).
25. Bode, W. & Schwager, P. *J. molec. Biol.* **98**, 693-717 (1975).
25. Delbaere, L. F. J., Hutcheon, W. L. B., James, M. N. G. & Thiessen, W. E. *Nature* **257**, 758-763 (1975).
27. Rossmann, M. G., Moras, D. & Olsen, K. W. *Nature* **250**, 194-199 (1974).
28. Eventoff, W. & Rossmann, M. G. *CRC Critical Reviews in Biochemistry* **11**, 111-140 (1974).
29. Schultz, G. E. & Schirmer, R. H. *Nature* **250**, 142-144 (1974).
30. Ohlsson, I., Nordström, B. & Brändén, C. I. *J. molec. Biol.* **89**, 339-354 (1974).
31. Holbrook, J. J., Liljas, A., Steindel, S. J. & Rossmann, M. G. *Enzymes* **11**, 191-292 (1975).
32. Watenpaugh, K. D., Sieker, L. C., Jensen, L. H., Le Gall, T. & Dubourdieu, M. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3185-3188 (1972).
33. Burnett, R. M. *et al.* *J. biol. Chem.* **249**, 4383-4392 (1974).
34. Olsen, K. W., Garavito, R. M., Saseban, M. N. & Rossmann, M. G. *J. molec. Biol.* **107**, 571-576 (1976).
35. Phillips, D. C., Rivers, P. S., Sternberg, M. J. E., Thornton, J. M. & Wilson, I. A. *Biochem. Soc. Trans.* **5**, 642-647 (1977).
36. Blow, D. M., Owen, M. G. & Nyburg, J. *Biochem. biophys. Res. Commun.* **67**, 728-734 (1977).
37. Drenth, J., Jansonius, J. N., Koekoek, R. & Wolthers, B. G. *Enzymes* **3**, 485-494 (1971).
38. Wright, C. S., Alden, R. A. & Kraut, J. *Nature* **221**, 235-242 (1969).
39. Drenth, J., Hol, W. G. J., Jansonius, J. N. & Koekoek, R. *Eur. J. Biochem.* **26**, 177-181 (1972).
40. Matthews, D. A., Alden, R. A., Birktoft, J. J., Freer, S. T. & Kraut, J. *J. biol. Chem.* **250**, 7120-7126 (1975).
41. Fersht, A. *Enzyme Structure and Mechanism* Ch. 10 (W. H. Freeman, San Francisco and Reading, 1978).

Binding of additional histones to chromatin core particles

Gerrit Voordouw* & Henryk Eisenberg

Polymer Department, The Weizmann Institute of Science, Rehovot, Israel

Isolated nucleosome core particles from chick erythrocytes and calf thymus have been used in experiments to show that chromatin subunits can absorb more than an additional equivalent of the histones making up the nucleosome octamer with no significant alteration in external shape and structure.

CHROMATIN is a complex of DNA, histone and non-histone proteins, the exact composition of which may be variable and is not accurately known. Grellet *et al.*¹ have, for instance, shown that the histone: DNA ratio of pea embryo axes decreases during germination. The ratio (w/w) in dry pea embryos, not yet germinated, was found to be approximately 3.3:1. These histones are not removed from the isolated chromatin by washing procedures or sucrose gradient centrifugation. As the germination of the embryos proceeds the histone: DNA ratio decreases to a lowest measured value of about 0.5:1. The decrease in histone:DNA ratio occurs in two steps and is

paralleled by a two-step increase in the total DNA content of the embryo and an increasing susceptibility of the isolated chromatin to nuclease digestion. These findings were interpreted by Grellet *et al.*¹ as a decrease in the number of nucleosomes along the DNA fibre as development, involving reactivation and transcription of the genome, proceeds. Although this interpretation may be correct it cannot explain the high histone: DNA ratios measured for the resting embryo. Accepting the common view²⁻⁴ of the nucleosome as an assembly of two of each of the histones H2A, H2B, H3 and H4 and minimally 140 base pairs of DNA, a maximal ratio of 1.2:1 can be calculated, which is much lower than the reported value of 3.3:1. The question remains, where and how are these additional histones bound? Are they bound in the nucleosome bead itself, do they bind to the spacer region, or is the whole structure fundamentally rearranged?

We were particularly interested in this question as we had found earlier that the highest histone:DNA ratio which we achieved by reconstituting complexes of these four histones and monodisperse ColE1-plasmid DNA was 2.4:1 (refs 5, 6). We therefore decided to examine the histone binding properties of isolated nucleosome core particles. This report presents studies of a preparation of chicken erythrocyte core particles, obtained

*Present address: Department of Biochemistry, Agricultural University, Wageningen, The Netherlands.

from J. F. Pardon, and our own preparation from calf thymus. The calf thymus preparation was found not to consist of pure core particles, yet the results obtained with it also support the idea that chromatin subunits may bind more than an additional equivalent of the histones constituting the nucleosome octamer without related gross changes in external shape and structure. Experiments were carried out in the absence of magnesium and of H1 histone.

Binding studies

Initial studies were carried out with calf-thymus chromatin particles prepared by digesting purified calf-thymus chromatin with staphylococcal nuclease, followed by rate zonal sucrose gradient centrifugation of the digest and precipitation with $MgCl_2$. Their histone:DNA ratio was found to be 1.3:1 (protein determined by the method of Lowry *et al.*⁷ using acid-extracted, H1-depleted calf-thymus histones as standards and DNA concentration from absorption measurements at 260 nm). Some properties of the purified particles were determined in a low ionic strength buffer (20 mM NaCl, 10 mM Tris HCl, pH 8, 0.1 mM EDTA). The chromatin particles sedimented as a single boundary in the analytical ultracentrifuge with a sedimentation coefficient $s_{23,s} = (11.9 \pm 0.1)S$ in the experimental conditions. When corrected to standard conditions, this gives $s_{20,w} = (11.2 \pm 0.1)S$. Their diffusion coefficient was measured both with the technique of laser light scattering and by combining the results of sedimentation velocity and sedimentation equilibrium studies (no value for the partial specific volume need be assumed for this analysis). Both methods gave values in good agreement and we obtained an average value, corrected to standard conditions, $D_{20,w} = (3.03 \pm 0.03) \times 10^{-7} cm^2 s^{-1}$. From the sedimentation equilibrium experiments a molecular weight $M_w = 265,000$ is calculated for these particles, when assuming⁸ a value $v = 0.661 g cm^{-3}$ for the partial specific volume. From the values $D_{20,w} = 3.9 \times 10^{-7}$ reported for core particle monomers⁸ and 2.16×10^{-7} for dimers⁹, we calculate that 3.03×10^{-7} could correspond to a mixture of mol fractions 0.8 and 0.2 of monomers and dimers, respectively. From the molecular weight $M_w = 201,000$ of a chromatin core particle of 140 base pairs of DNA and 8 histones we calculate $M_w = 270,000$ for our particles, consistent with the value obtained from sedimentation equilibrium. Another possibility is that the particles described here are similar to those of the rat liver mononucleosomes studied by Sanders and Hsu¹⁰.

When a mixture of the histones H2A, H2B, H3 and H4, obtained by acid extraction of calf-thymus tissue¹¹ and purification by Biogel P-60 gel-chromatography¹² was added to a solution of calf-thymus chromatin particles at low ionic strength an immediate precipitation of the particles was observed. However, at high ionic strength (0.5 M NaCl, 1 mM Tris HCl, 0.1 mM EDTA, pH 8) they stayed in solution. We therefore decided to titrate the particles with the histone mixture at the higher ionic strength, following the course of the titration by measuring $s_{23,s}$ (the sedimentation coefficient at 23 °C in the stated experimental conditions). The results of this titration experiment are shown in Fig. 1. When no extra histones are added to the solution a value $s_{23,s} = (10.4 \pm 0.1)S$ is found for the particles. Addition of increasing amounts of the histone mixture results in an increasing $s_{23,s}$ until a 1.2–1.5-fold excess of added histone DNA has been reached. Further increase in the amount of added histone does not result in a further increase of $s_{23,s}$. Its limiting value is $s_{23,s} = (12.5 \pm 0.1)S$. Provided the histones are bound stoichiometrically in the presaturation range it follows that the actual ratio of bound histones DNA is about 2.5:1 in the plateau region. Stoichiometric binding could be demonstrated conclusively as follows. The calf-thymus chromatin particles were mixed with a 4:1 excess of added histones–DNA and the resulting complex was purified from unbound histones by sucrose gradient centrifugation. After dialysis the histone:DNA ratio was found to be 2.5:1. The sedimentation coefficients of the complex before and after sucrose gradient centrifugation were identical.

We therefore conclude that the added histones bind stoichiometrically up to the saturation point. The binding must be quite tight in 0.5 M NaCl, as the complex can be isolated by sucrose gradient centrifugation, apparently without dissociation of bound histones.

Effect of histone binding on particle size

For the diffusion coefficient of the complex saturated with bound histones in the 0.5 M NaCl buffer we calculate $D_{20,w} = (3.13 \pm 0.06) \times 10^{-7} cm^2 s^{-1}$ from a combination of sedimentation equilibrium and velocity results. This value is not significantly different from that obtained for the particle in 20 mM NaCl buffer. We conclude, therefore, that the binding of additional histones does not lead to a significant change in the hydrodynamic size of these particles. However, their hydrodynamic size, derived from the diffusion coefficient, is larger than that of the core particle. The comparison is more easily made in terms of Stokes' radius, which is 68–70 Å for the particles studied here and 55 Å for core particles examined by Olins *et al.*⁸.

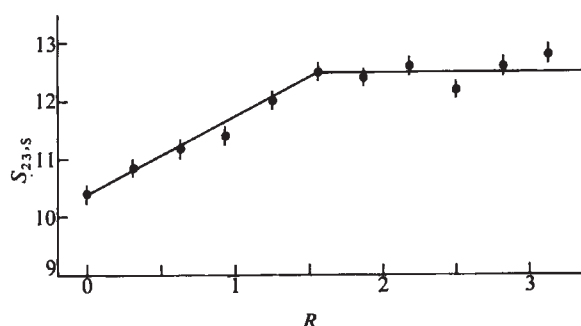


Fig. 1 Titration of calf-thymus chromatin particles with an approximately equimolar mixture of the histones H2A, H2B, H3 and H4. The ordinate represents the sedimentation coefficient, $s_{23,s}$ observed in the experimental conditions (0.5 M NaCl, 1 mM Tris-HCl, pH 8, 0.1 mM EDTA). The abscissa represents the ratio (w/w) of histones added in excess of DNA. For this experiment 20–50 μ l of a stock solution of particles ($A_{260} = 7.5$ in 0.2 mM EDTA pH 8), was diluted into 1 ml of a solution containing the desired amount of histones in the 0.5M NaCl solvent. After mixing the sedimentation velocities of the particle-histone complexes formed were determined using a Beckman Model E analytical ultracentrifuge, equipped with electronic speed control, a photoelectric scanner system and multiplex accessory unit. The scanning wavelength was 265 nm. Cells fitted with a 12-mm filled Epon double sector centrepiece placed in either a two-hole AnH or a four-hole AnF titanium rotor were used for the runs. Rotor speeds were either 36,000 or 40,000 r.p.m. All runs were done at room temperature (22–25 °C) without use of the RTIC temperature regulating system. Temperatures were constant to within 0.2 °C for a given run. The measured sedimentation coefficients, $s_{t,s}$, were corrected to their value at 23 °C, $s_{23,s}$, as described elsewhere⁸.

The DNA of the chicken erythrocyte core particles moved as a single band on 3.5% polyacrylamide gels for double-stranded DNA and on Maniatis gels¹³ for single stranded DNA. The length of the DNA was estimated as 145 ± 5 base pairs by comparison with the DNA from a DNase I digest of chromatin which contains a series of bands differing by 10 bases. The histone: DNA ratio was 1:1 as determined by the method described above. This is somewhat lower than the theoretical value for a core particle. At low ionic strength a value $s_{20,w} = 11.5 S$ was obtained for the sedimentation coefficient, corrected to standard conditions. Sedimentation equilibrium experiments gave a molecular weight $M_w = 196,000 \pm 10,000$. For the diffusion coefficient corrected to standard conditions we calculate $D_{20,w} = (4.2 \pm 0.2) \times 10^{-7} cm^2 s^{-1}$ from the combined results of our sedimentation equilibrium and velocity studies. From this value we obtain a Stokes' radius $R_s = 51 \text{ Å}$, which is consistent with the dimensions of core particles determined by X-ray diffraction and X-ray and neutron scattering

studies¹⁴⁻¹⁷. These have shown the core particle to have the shape of a flat cylinder with a diameter of 110 Å and a height of 57 Å.

The addition of progressive amounts of calf-thymus histones to a solution of chicken erythrocyte chromatin core particles at high ionic strength (0.5 M NaCl, 1 mM Tris-HCl, pH 8, 0.1 mM EDTA) was again monitored with the sedimentation velocity technique (Fig. 2). Without addition of histones the core particles sedimented with $s_{23,s} = (11.0 \pm 0.1)S$. Addition of histones raises the value for $s_{23,s}$ observed in the experimental conditions to a limiting value of $s_{23,s} = (13.5 \pm 0.1)S$. The limiting value is reached when a 2.0–2.5-fold excess of calf-thymus histones:DNA has been added to the core particle solution. Mixing the core particle solution with a 4:1 excess of

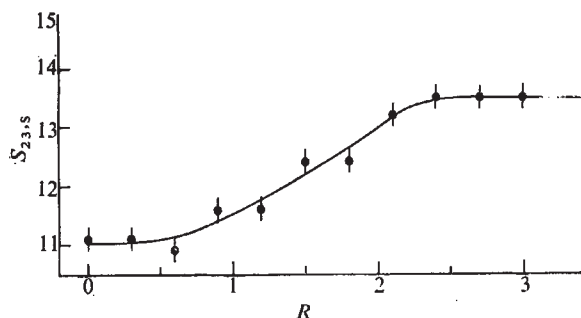


Fig. 2 Titration of chicken erythrocyte core particles with an approximately equimolar mixture of the histones H2A, H2B, H3 and H4 from calf-thymus. Symbols and experimental procedure were as for Fig. 1, except that all results were obtained using a 20 mm centrepiece and an MSE analytical ultracentrifuge.

added histones–DNA and purifying the resulting complex by sedimentation through a 5–10% (w/w) gradient of sucrose in 0.5 M NaCl, 1 mM Tris HCl, pH 8, 0.1 mM EDTA gave a particle with a histone:DNA ratio of 2.9:1. We thus conclude that, as in the case of the calf-thymus chromatin particles, the histones were bound stoichiometrically in the presaturation range. Analysis of the sedimentation equilibrium and sedimentation velocity data for this histone saturated particle gives $D_{20,w} = 3.51 \pm 0.08 \times 10^{-7} \text{ cm}^2 \text{ s}^{-1}$, and hence a Stokes' radius $R_s = 61 \text{ Å}$. The hydrodynamic dimensions of the chicken erythrocyte core particle thus expand somewhat on binding of additional histones. We emphasise that this is not necessarily to be expected. Accepting the view of the nucleosome core as a flat cylindrical particle^{16,17}, we calculate a volume of 542 nm³ per particle. It is easily estimated that the volume of the components of the histone octamer is about 132 nm³ and that of 140 base pairs of DNA about 84 nm³. The total volume occupied in the core particle by closely packed protein and nucleic acid is then only 216 nm³, leaving ample space (326 nm³) for water or additional histone or non-histone components. We calculate about 1 g water per g particles if all free space is occupied by water. This is considerably in excess of a conservative estimate¹⁸ of 0.2 to 0.3 g water per g of nucleic acid or protein in usual situations.

The excess of histones could, therefore, in principle be bound within the original hydrodynamic domain of the core particle. However, binding outside this domain is also possible. Finch *et al.*¹⁶ describe the cylindrical core particle as being divided into two layers of height 28 Å. Stacking of another two of these layers perpendicular to the cylinder axis would lead to a cylindrical particle of height 114 Å and diameter 110 Å, with a histone: DNA ratio of 2.4:1. This particle can be approximated by a sphere with a radius between 57 and 77 Å, while the Stokes' radius calculated for the histone-saturated core particles is 61 Å. It is obvious that the question

of where these additionally bound histones are located in the core particle structure will only be answered after more detailed structural investigations on a number of carefully selected systems.

Discussion

Thus we have shown that the histone: DNA ratio of chromatin core particles can be raised by more than twofold over the theoretical value for a 140-base pair piece of DNA and a histone octamer (1.2:1). This satisfies at least in part, the objectives stated at the beginning of this article. Other workers^{19,20} have shown that presumably a tetramer of H3 and H4 alone can fold the DNA in a nucleosome-like structure. The question then arises as to whether it is still adequate to consider the nucleosome exclusively as an assembly of eight histones and a 140–200-base pair piece of DNA. It is more likely that there is a spectrum of particles along the DNA with histone: DNA ratios ranging from 0.6:1 up to 3.1. However, in the conditions of chromatin digestion (low ionic strength) the particles or structures with a low histone: DNA ratio are preferentially degraded, while those with high histone: DNA ratios are not soluble. The result of a nuclease digestion at low ionic strength may thus be a collection of particles with an average of about eight histones per particle. A distribution of histone to DNA ratios is indicated, however, by the fact that preparations of core particles can generally be separated into several subfractions on the basis of solubility and electrophoretic phenomena (for example, see ref. 8).

Finally, the binding of additional histones to chromatin core particles seems to be a spontaneous process with a large decrease in the free energy content of the system. The presence of a high histone: DNA ratio along the chromatin fibre must further stabilise the local histone: DNA interactions in the complex and thus prevent the DNA from being either replicated or transcribed insofar as these processes depend on interactions of specific proteins with stretches of naked DNA. The reason for the high histone: DNA ratios in genetically dormant systems such as resting seeds¹ or fully differentiated cells²¹ is understandable following the present interpretation.

This study was supported in part by a grant from the U.S.–Israel Binational Science Foundation, Jerusalem, Israel. We thank Drs M. A. Hirshfeld, Z. Kam, T. Vogel and H. W. J. van den Broek for useful discussions, Mr T. Bettecken and Mrs N. Borochoy for assistance in some of the measurements, and Dr J. F. Pardon and Mrs Myra Jacobs for providing a sample of chicken erythrocyte chromatin core particles and for carrying out the gel electrophoretic characterisation of this sample. G.V. acknowledges the support of an EMBO short-term fellowship.

Received 22 February; accepted 31 March 1978.

- Grellet, F., Delseney, M. & Guitton, Y. *Nature* **267**, 724–726 (1977).
- Söllner-Webb, B. & Felsenfeld, G. *Biochemistry* **14**, 2915–2920 (1975).
- Shaw, B. R., Herman, T. M., Kovacic, R. T., Beaudreau, G. S. & Van Holde, K. E. *Proc. natn. Acad. Sci. U.S.A.* **73**, 505–509 (1976).
- Kornberg, R. D. *Science* **184**, 868–871 (1974).
- Voordouw, G., Kalif, D. & Eisenberg, H. *Nucleic Acids Res.* **4**, 1207–1223 (1977).
- Voordouw, G., Kam, Z., Borochoy, N. & Eisenberg, H. (in preparation).
- Lowry, D. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
- Olins, A. L., Carlson, H. D., Wright, E. B. & Olins, D. E. *Nucleic Acids Res.* **3**, 3271–3291 (1976).
- Schmitz, K. S. & Shaw, B. R. *Science* **197**, 661–663 (1977); *Biopolymers* **16**, 2619–2633 (1977).
- Sanders, M. M. & Hsu, J. T. *Biochemistry* **16**, 1690–1695 (1977).
- Kobayashi, Y. & Iwai, K. *J. Biochem.* **67**, 465–472 (1970).
- Böhm, E. L. *et al. FEBS Lett.* **34**, 217–221 (1973).
- Maniatis, T., Jeffrey, A. & Van de Sande, H. *Biochemistry* **14**, 3787–3794 (1975).
- Pardon, J. F. *et al. Nucleic Acids Res.* **2**, 2163–2176 (1975).
- Hjelm, P., Kneale, G. G., Suau, P., Baldwin, J. P. & Bradbury, E. M. *Cell* **10**, 139–151 (1977).
- Finch, J. T. *et al. Nature* **269**, 29–36 (1977).
- Pardon, J. F. *et al. Nucleic Acids Res.* **4**, 3199–3214 (1977).
- Reisler, E., Haik, Y. & Eisenberg, H. *Biochemistry* **16**, 197–203 (1977).
- Camerini-Otero, P. D., Söllner-Webb, B. & Felsenfeld, G. *Cell* **8**, 333–347 (1976).
- Söllner-Webb, Camerini-Otero, R. D. & Felsenfeld, G. *Cell* **9**, 179–193 (1976).
- Neelin, J. M., Mazen, A. & Champagne, M. *FEBS Lett.* **65**, 309–314 (1976).

letters to nature

The present appearance of white holes

It has been argued¹ that even though a stable model of a white hole requires essentially zero delay between the big bang and the inception of expansion of the white hole, one can (at least in principle) still observe certain types of white hole, most likely those which never expand beyond their 'Schwarzschild spheres' into our space, and which have, in fact, always been visible. Here we report the main results of an analysis of the optical properties of such white holes, the complete details of which will be published elsewhere.

Previous discussions of the optical appearance of white holes have been concerned with the analysis of only radial null geodesics and have assumed that radiation from white holes is strongly blueshifted^{2,3}. We find that radiation from the white holes mentioned above is strongly blueshifted for only an insignificant part of their early histories (K.L. and R.C.R., in preparation). Moreover, our calculations show that nonradial photons completely determine the optical appearance of the

Fig. 1 The distributions of redshift and surface brightness (assuming constant luminosity) as functions of impact parameter seen by a distant observer when the spherical, expanding object with $\gamma=0$ has a radius $r/m=1$.

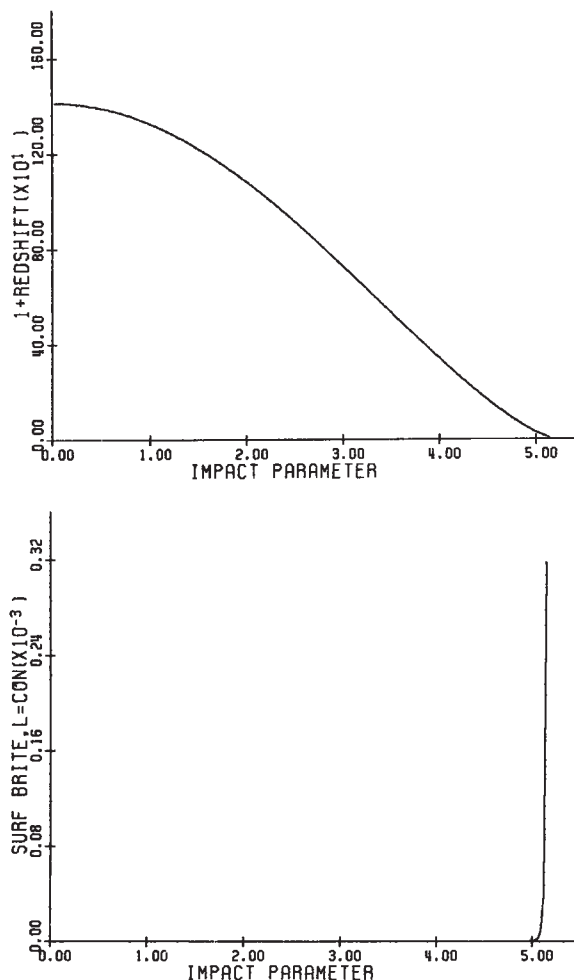
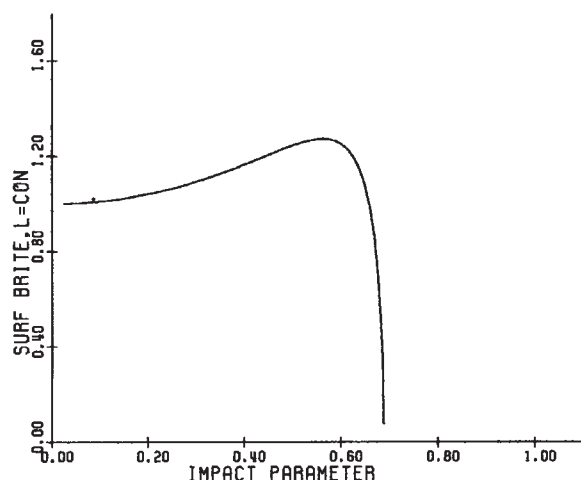
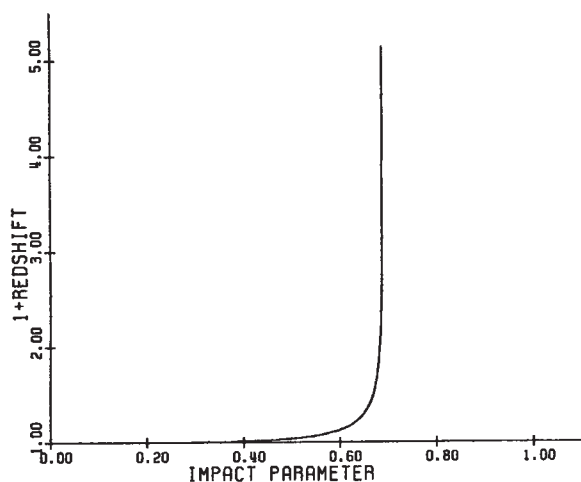


Fig. 2 The same as Fig. 1, but with the radius now $r/m=2 \times 10^{-6}$. Note that the redshift scale is to be multiplied by 10, while the surface brightness scale is to be multiplied by 10^{-3} . The maximum value of $(1+z)$ is 1,412 and the minimum value 3.80.

white hole at late times. With the aid of hindsight this latter result is not surprising in view of the optical characteristics of black holes^{4,5}. Qualitatively, our main conclusions are the following. (1) The early history of such a white hole determines its appearance at late observer times; and (2), radiation from such an object is mostly redshifted.

We have considered a white hole in the Schwarzschild spacetime and a distant static observer in an asymptotically flat region. The photons which are emitted simultaneously on the surface of the white hole will not reach the observer simultaneously. Rather, those photons which do reach the observer simultaneously determine a trajectory (locus of emission points) in the spacetime history of the white hole's surface. This trajectory is parameterised by the dimensionless photon impact parameter b/m ($b=0$ for the radial photon which appears at the centre of the observer's disk). Numerical calculation of the trajectory yields the optical characteristics of the object at an instant of the observer's time.

We give here some typical results of such a calculation for the case of a 'grey hole'⁶ (the case $\gamma = 0$ of ref. 1). In Figs 1 and 2 we show the distribution of frequency shifts $(1+z)$ across the observers entire optical disk, and the observed surface brightness distribution under the assumption of isotropic emission in the rest frame on the white hole. Figure 1 shows the appearance when the white hole has expanded to half its Schwarzschild radius, whereas for Fig. 2 the white hole has expanded to $(1-5 \times 10^{-7})$ of its Schwarzschild radius. Even though the white hole has expanded almost to its Schwarzschild radius, it can be shown that the time measured by the external observer since the inception of expansion is only $\sim 27GM/c^3$, or for $M \sim 10^{12}M_{\odot}$, only about 4 yr! Note that the 'white' hole already appears as a highly redshifted and sharply peaked ring. To the distant observer this ring subtends such a small half angle that the 'white' hole in fact appears as a dim, red, unresolved point source.

In conclusion, we regret to report that our calculations show that one cannot invoke the concept of a simple, stable, white hole as a natural explanation of highly energetic astrophysical phenomena. If any white hole were to serve directly as a source of even moderately energetic photons, it would have to be the classical type of object which enters our section of the Schwarzschild spacetime with $\gamma \rightarrow 1$. The problem then is that if such an object has a long expansion timescale, it is at present much too large to serve as a compact source of energy required, say, for certain types of galactic nuclei.

We thank Dr John Roth (University of Toronto Computing Centre) for help with the computer program, Drs H. C. Corben, C. C. Dyer, and E. Honig for constructive comments, and the NRC of Canada for financial support.

KAYLL LAKE*

R. C. ROEDER

Scarborough College and David Dunlap Observatory,
University of Toronto,
Toronto, Ontario, Canada

*Present address: Theoretical Physics Institute, University of Alberta,
Edmonton, Alberta, Canada.

Received 8 March; accepted 29 March 1978.

1. Lake, K. *Nature* 272, 599-601 (1978).
2. Narlikar, J. V., Apparao, K. & Dadhich, N. *Nature* 251, 590 (1974).
3. Narlikar, J. V. & Apparao, K. *Astrophys. Space Sci.* 35, 321 (1975).
4. Ames, W. & Thorne, K. S. *Astrophys. J.* 151, 659 (1968).
5. Jaffe, J. *Ann. Phys.* 55, 374 (1969).
6. Novikov, I. D. & Thorne, K. S. in *Black Holes* (Gordon and Breach, New York, 1973).

4U0241+61: a luminous low-redshift QSO

A PRECISE (30'') celestial position for the faint (1.8 μ Jy, 2-11 keV) X-ray source 4U0241+61 (ref. 1) and its identification at optical and radio wavelengths are reported here. Despite its low galactic latitude ($b'' = +2.4$), the source is extragalactic: a luminous ($M_v = 25.8 \pm 1.8$), low-redshift ($z = 0.0438$) QSO with relatively intense radio emission (0.2-0.5 Jy).

The X-ray observation was motivated by the fact that the 0.4 deg² error region of 4U0241+61 overlaps the 3 deg² error region of the COS B γ -ray source CG135+1 (ref. 2). The observation was made with the rotation modulation collimator detectors³⁻⁵ on the SAS 3 satellite. Data from 82 orbits, obtained 23-30 November 1977, provided 2.0×10^6 s of effective exposure. The source 4U0241+61 was detected with each of the two detection systems (MC5 and MC9) at a significance of $\sim 10 \sigma$. These data yielded two quasi-independent positions separated by 20''. The coordinates for the average position (designated 2S0241+622) are α (1950) = 2 h 41 min 01.3 s δ (1950) = 62°15'27'' with a 90% confidence error radius of 30''. Figure 1 shows this position in relation to the error

regions of 4U0241+61 and CG135+1. The finding chart in Fig. 2 shows the SAS 3 error circle with the QSO indicated.

The X-ray counting rates for the two energy channels in the two detectors are given in Table 1. These counting rates are consistent with the intensity obtained by applying a point summation technique to the Uhuru data base⁶. We cannot determine a unique X-ray spectrum from the rates in two energy channels. If we assume a low-energy cutoff below 2 keV, the rates are consistent with a power law energy spectrum of spectral index 0.4 ± 0.3 . The 2-11 keV flux is 4.0×10^{-11} erg cm⁻² s⁻¹. At the redshift distance of 240 Mpc ($H_0 = 55$ km s⁻¹ Mpc⁻¹) this gives an X-ray luminosity of 2.7×10^{44} erg s⁻¹.

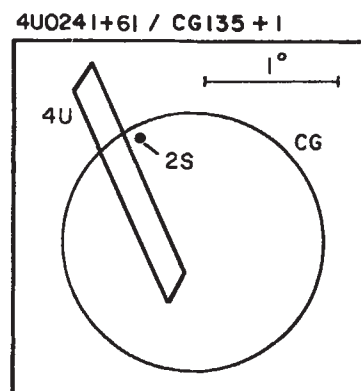


Fig. 1 SAS 3 position (2S) for the X-ray source 4U0214+61, the Uhuru position¹ (4U) for that source, and the COS B position² (CG) for the γ -ray source CG135+1.

The ratio of counting rates in the two channels is comparable to that measured with the same detectors for the X-ray QSO MR2251-178 (ref. 7) and for the Seyfert Type 1 galaxy 3C120 (ref. 8).

Spectrophotometry of the brightest object in the field of the SAS 3 position was obtained (by B.M.) on 4 December 1977 with the Robinson-Wampler image tube scanner⁹ at the Cassegrain focus of the Lick Observatory 3-m Shane reflector. The grating used provides approximately 8 Å resolution throughout the 3,700-7,900 Å range. Absolute calibration

2S0241+622

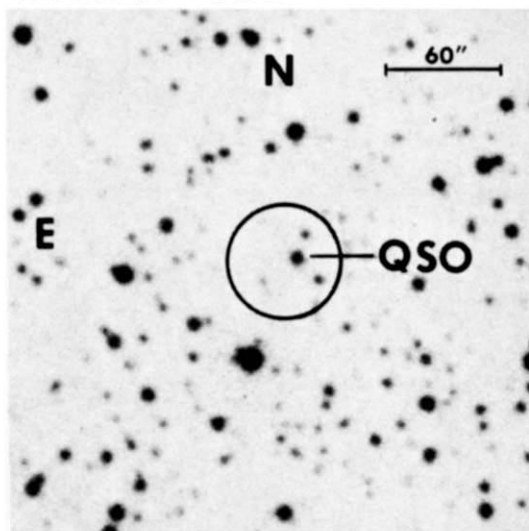


Fig. 2 Finding chart for the SAS 3 position of 4U0241+61. QSO0241+622 is indicated. Chart from the Palomar Observatory Sky Survey (National Geographic Society) E (red) print.

Table 1 X-ray and radio flux densities* for 0241+622

X-ray energy	MC5	Counting rate MC9	Average	Flux density†
2-6 keV	$0.42 \pm 0.05 \text{ s}^{-1}$	$0.38 \pm 0.04 \text{ s}^{-1}$	$0.40 \pm 0.03 \text{ s}^{-1}$	$2.2 \pm 0.2 \text{ } \mu\text{Jy}$
6-11 keV	0.14 ± 0.03	0.18 ± 0.03	0.16 ± 0.02	1.6 ± 0.2
Radio frequency	5 Dec 1977	17 Dec 1977	Average	
1465 MHz	$0.27 \pm 0.02 \text{ Jy}$	$0.31 \pm 0.02 \text{ Jy}$	$0.29 \pm 0.02 \text{ Jy}$	
4885	0.20 ± 0.01	0.19 ± 0.01	0.20 ± 0.01	
14765	0.48 ± 0.05	0.47 ± 0.05	0.48 ± 0.05	
22485	0.53 ± 0.05	0.53 ± 0.05	0.53 ± 0.05	

*Errors quoted are measurement uncertainties.

†For a power law energy spectrum with index 0.4; $1.0 \text{ } \mu\text{Jy} = 0.24 \times 10^{-11} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ keV}^{-1}$.

was obtained by observing standard stars¹⁰. Less than one minute of observation was required to establish that this object is a redshifted, extragalactic, emission-line object. Several scans, totalling 72 min of exposure, yielded a spectrum with good statistics between 4,300 Å and 7,700 Å. Heavy interstellar extinction severely depresses the continuum short of 4,300 Å.

The spectrum is dominated by very broad and intense H α emission, which reaches $\sim 250\%$ of the continuum strength. Much weaker, broad H β and H γ emission lines are also visible. Strong narrow emission due to [OIII] $\lambda\lambda 4,959, 5,007$ is present, as well as barely detectable but definitely present H ϵ $\lambda 5,876$ and [NIII] $\lambda\lambda 3,869, 3,968$. The forbidden lines yield a redshift $z = 0.0438 \pm 0.0001$. The Balmer lines are consistent with this redshift, but are too broad to provide accurate independent measures. The spectrum also shows very prominent, unredshifted interstellar absorption lines and bands—NaD $\lambda 4,430$, $\lambda 5,778$ – $5,780$, and $\lambda 6,284$, consistent with the heavy obscuration in this region.

The morphology and luminosity of 0241+622 meet the classical definitions of a QSO. The object is apparently stellar on the blue plate of the Palomar Observatory Sky Survey. The red image seems slightly asymmetric, although not convincingly extended. At the Cassegrain acquisition television of the Shane telescope, which is red sensitive, the object seemed to be stellar in moderately good seeing conditions ($\sim 1''$). A direct plate ($40'' \text{ mm}^{-1}$) obtained with the Lick 0.9-m Crossley reflector on 13 December 1977 also shows a stellar image. The spectral response of the 098 emulsion that was used overlaps the Sky Survey red bandpass. Thus any extended structure underlying the bright nucleus must be of moderately low surface brightness.

A calculation of the absolute visual magnitude requires a correction for interstellar extinction. We have estimated this correction by measuring the equivalent widths of the interstellar $\lambda 4,430$ band and (unresolved) $\lambda 5,778$ – $5,780$ band. Both imply¹¹ a colour excess of $E(B-V) = 1.4 \pm 0.5 \text{ mag}$. The most prominent absorption feature in the spectrum, NaD, generally shows less correlation with colour excess and, thus, was not used. Accurate photometry of this object is not yet available. The diameter of the Sky Survey images indicate^{12,13} $B = 16.7 \pm 0.5 \text{ mag}$, $R = 15.3 \pm 0.5 \text{ mag}$. The Crossley plate described above, obtained immediately after the X-ray, optical, and radio observations, shows that the current brightness is comparable to that on the Sky Survey print. An appropriate colour equation¹² yields $V \sim 15.7 \pm 0.5$. The spectrophotometric data imply a somewhat redder colour, $(B-V) \sim 1.4$. If we adopt $A_V/E(B-V) = 3.3$, we derive $A_V = 4.6 \pm 1.7 \text{ mag}$, and $M_V = -25.8 \pm 1.8$. This is near the mean luminosity of the QSOs tabulated by Sandage¹⁴ and significantly brighter ($\sim 3 \text{ mag}$) than the most luminous Seyfert Type I measured by Weedman¹⁵.

The compilation of 633 QSOs by Burbidge *et al.* lists only one object, 1301+375 (= B272) at a lower redshift¹⁷, $z = 0.036$. However, this object has $M_V \approx -19$ and is probably a compact subluminal galaxy. Thus, 0241+622 is the closest known

QSO. If this luminous QSO were fortuitously located outside of the galactic plane where extinction is negligible, it would be an $\sim 11 \text{ mag}$ object.

Radio observations of the X-ray source region (by R.H.) were carried out on 5 December and 17 December 1977 with seven antennae of the very large array (VLA) now under construction in New Mexico. A strong point source was found at all four VLA frequencies located at the position: $\alpha(1950) = 2 \text{ h } 41 \text{ min } 00.68 \text{ s} \pm 0.03 \text{ s}$; $\delta(1950) = 62^\circ 15' 27.5'' \pm 0.5''$. This is coincident with the optical object to within $2''$. The observed radio flux densities are summarised in Table 1. No significant flux differences are apparent from the two days of observations. The flux density calibration is based upon assumed flux densities of 1.04, 1.00, 1.5, and 1.7 Jy at 1,465, 4,885, 14,765, and 22,485 MHz respectively, for the calibrator D0224+67. The radio spectrum is typical of multi-component synchrotron spectra associated with quasars or galactic nuclei in which more recent synchrotron events add optically thick components to optically thin non-thermal spectra due to older events.

The source was also observed with the beam switching system of the Effelsberg 100-m dish at 22,268 MHz on 20 December 1977. The flux density was $0.45 \pm 0.04 \text{ Jy}$ (D. Downes and K. Weiler, personal communication). A flux density of 6.0 Jy was assumed for the comparison source NGC7027.

Fig. 3 Flux densities for the three X-ray QSOs. References are given in Table 2. The measurement errors for the present work (0241+622) are indicated. The optical data points are corrected for interstellar extinction. The data now available for 0241+622 are consistent with constant intensities. The variabilities for the other sources, where known, range up to a factor of 8 and are given in Table 2. The optical fluxes are the brightest among all known QSOs. X-ray fluxes are near the sensitivity limits of present experiments. $1.0 \text{ Uhuru count s}^{-1} = 1.7 \times 10^{-8} \text{ Jy}$ (2-6 keV).

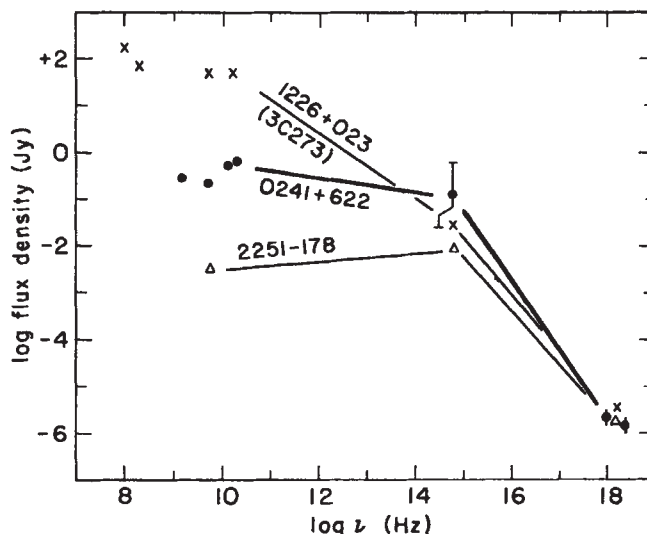


Table 2 Properties of X-ray emitting QSOs and Seyferts

Source	Other names	Redshift	S_{ν} (5 GHz) Jy ($\text{erg s}^{-1} \text{GHz}^{-1}$)	M_{ν}^*	L_x^* (2–11 keV) $10^{44} \text{ erg s}^{-1}$	Ref.
0241+622	4U0241+61 2S0241+622	0.044	0.2 (1.4×10^{40})	$-25.8 \pm 1.8^\dagger$	3	See text
1226+023	3C273 4U1226+02 2A1225+022	0.158	36–44 ‡ (3.6×10^{43})	-27.0 to $-27.7^\ddagger$	20–100 ‡	1, 16, 22, 31–33
2251–178	2A2251–179 MR2251–178 2S2251–178	0.064	0.003 (4.4×10^{38})	-23.9	2–14 ‡	7, 22, 23
0121–590 (Spiral+QSO)	ESO113–IG45 Fairall 9 2A0120–591 4U0106–59	0.045		-24.0	2	1, 22, 27, 28, 38
X-ray Seyferts	Minimum: Maximum:	0.003 0.090	<0.01 6.0	-19.0 -23.8	0.03 13	24–26, 31, 34–37

*For $H_0 = 55 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

† Uncertainty dominated by correction for interstellar extinction.

‡ Reported variation in intensity.

The VLA data at 22,485 MHz and baseline (D/λ) of 7.89×10^5 yield an upper limit of $0.1''$ for the angular extent of 0241+622. Upper limits at other frequencies and baselines are $0.2''$ (14,765 MHz, $D/\lambda = 5.17 \times 10^5$), $0.5''$ (4,885 MHz, $D/\lambda = 1.71 \times 10^5$), and $1.8''$ (1,465 MHz, $D/\lambda = 5.13 \times 10^4$). The upper limit of $0.1''$ at a distance of 240 Mpc gives an upper limit of 120 pc for the source size. Assuming this size limit applies at 1,465 MHz, we obtain a lower limit of $2.4 \times 10^7 \text{ K}$ for the brightness temperature of the source. These size and temperature limits support the conclusion that the source is a QSO.

The QSO 0241+622 is located within the 3 deg^2 error region of the COS B (γ -ray source CG135+1 (refs 2, 18). The association of the γ -ray and X-ray sources is obviously speculative because of the large size of the γ -ray error region. The extragalactic nature of 4U0241+61 argues against the association as all the COS B sources lie close to the galactic plane and, thus, are probably of galactic origin². Also, we note that a highly variable radio source¹⁹, GT0236+610, associated with an OB star^{20,21} has been discovered within the CG135+1 error region. The present SAS 3 data yield an upper limit (95% confidence) of $0.4 \mu\text{Jy}$ (2–11 keV) for an X-ray source at this position. This radio source is $1.3''$ from 0241+622.

The QSO 0241+622 is the third positively identified X-ray emitting QSO. 3C273 has been convincingly associated with the X-ray source 4U1226+02 on the basis of the Uhuru¹ and SAS 3 (H. Schnopper, personal communication) positions. The X-ray source 2A2251–179 (ref. 22) has been identified recently with a previously unrecorded QSO (MR2251–178) on the basis of a SAS 3 position⁷. The energy spectra of the three X-ray QSOs are shown in Fig. 3. The apparent optical fluxes are the highest among all known QSOs¹⁸ ($m_v = 12.9$ for 3C273 and 14.1 for MR2251–178; ref. 23). The X-ray fluxes all lie near the detection thresholds of present experiments. Deeper X-ray surveys with good positional precision will probably lead to the detection of emission from known fainter QSOs and the discovery of many presently unrecorded QSOs.

Table 2 summarises properties of the X-ray QSOs and, for comparison, the range of properties of X-ray emitting Seyfert galaxies^{24–26}. Also listed are the properties of Fairall 9 – ESO113–IG45, a spiral galaxy with a QSO-like nucleus, which is probably associated^{27,28} with 2A0120–591 (.... 4U0106–59) (refs 1, 22). The radio and optical properties of many Seyferts and QSOs show considerable overlap²⁹, and this has been considered as evidence that the two classes are closely related³⁰. As shown in Table 2, the X-ray luminosities of X-ray emitting Seyferts and QSOs also overlap. This is further evidence for the close relationship of these objects.

We thank the staffs of the Center for Space Research at MIT and Goddard Space Flight Center of NASA for support,

also G. Grueff, W. Lewin, G. Ricker, & B. Swanenberg for helpful conversations and E. A. Harlan and K. Witter for aid in obtaining the spectroscopy and photography. This work was supported in part by NASA under contract NAS5-11450. B.M. acknowledges the financial support of the US NSF. The National Radio Astronomy Observatory is operated by Associated Universities, Incorporated under contract with the US NSF.

K. M. V. APPARAO

Tata Institute of Fundamental Research,
Bombay, India

G. F. BIGNAMI

L. MARASCHI

Laboratorio di Fisica Cosmica e Tecnologie Relative,
CNR, Milan, Italy

H. HELMKEN

Harvard-Smithsonian Center for Astrophysics,
Cambridge, Massachusetts 02138

B. MARGON

Department of Astronomy, University of California,
Los Angeles, California 90024

R. HJELLMING

National Radio Astronomy Observatory, VLA Project,
Socorro, New Mexico 87801

H. V. BRADT

R. G. DOWER

Center for Space Research and Department of Physics,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

Received 20 March; accepted 11 April 1978.

- Forman, W. *et al. Astrophys. J. Suppl.* (in the press).
- Hermesen, W. *et al. Nature* 269, 494–495 (1977).
- Doxsey, R. *et al. Astrophys. J. Lett.* 203, L9–L12 (1976).
- Schnopper, H. W. *et al. Astrophys. J. Lett.* 210, L75–L77 (1976).
- Doxsey, R. E., Apparao, K. M. V., Bradt, H. V., Dower, R. G. & Jernigan, J. G. *Nature* 269, 112–116 (1977).
- Julien, P. & Helmken, H. *Nature* 272, 699–701 (1978).
- Ricker, G. R. *et al. Nature* 271, 35–37 (1978).
- Schnopper, H. W. *et al. Astrophys. J. Lett.* 215, L7–L11 (1977).
- Robinson, L. B. & Wampler, E. J. *Publ. astr. Soc. Pac.* 84, 161–166 (1972).
- Stone, R. P. S. *Astrophys. J.* 218, 767–769 (1977).
- Herbig, G. H. *Astrophys. J.* 196, 129–160 (1975).
- King, I. R. & Raffi, M. I. *Publ. astr. Soc. Pac.* 89, 120–121 (1977).
- Liller, M. H. & Liller, W. *Astrophys. J. Lett.* 199, L133–L135 (1975).
- Weedman, D. W. *Astrophys. J.* 178, 25–44 (1972).
- Sandage, A. *Astrophys. J.* 183, 29–39 (1973).
- Burbidge, G. R., Crowne, A. H. & Smith, H. E. *Astrophys. J. Suppl.* 33, 113–188 (1977).
- Braccetti, A., Formigini, L. & Gandolfi, E. *Astr. Astrophys.* 5, 264–279 (1970).
- Maraschi, L. *et al. Nature* 272, 679–681 (1978).
- Gregory, P. C. & Taylor, R. *IAU Circ.* No. 3164 (1978).
- Sanduleak, N. *IAU Circ.* No. 3170 (1978).

21. Hjellming, R., Hogg, D., Hvatum, H., Gregory, P. & Taylor, R. *IAU Circ.* No. 3170 (1978).
22. Cooke, B. A. *et al. Mon. Not. R. astr. Soc.* 182, 489–515 (1978).
23. Canizares, C. R., McClintock, J. E. & Ricker, G. R. *Astrophys. J. Lett.* (submitted).
24. Elvis, M. *et al. Mon. Not. R. astr. Soc.* 183, 129–681 (1978).
25. Tananbaum, H. *et al. Astrophys. J. Lett.* (in the press).
26. Schnopper, H. W. & Cash, W. *IAU Circ.* No. 3154 (1977).
27. Ricker, G. R. *Nature* 271, 334 (1978).
28. West, R. M., Danks, A. C. & Alcaïno, G. *Astr. Astrophys.* 62, L13–L16 (1978).
29. Weedman, D. W. *Astrophys. J.* 208, 30–36 (1976).
30. Weedman, D. W. *Q. Jl R. astr. Soc.* 17, 227–262 (1976).
31. Kellermann, K. I. & Pauliny-Toth, I. I. K. *Astrophys. Lett.* 8, 153–160 (1971).
32. Dent, W. A. & Kapitzky, J. E. *Astr. J.* 81, 1053–1068 (1976).
33. Smith, H. J. & Hoffleit, D. *Nature* 198, 650–651 (1963).
34. de Bruyn, A. G. & Wilson, A. S. *Astr. Astrophys.* 53, 93–105 (1976).
35. Wright, A. E., Allen, D. A., Krug, P. A., Morton, D. C. & Smith, M. G. *IAU Circ.* No. 3145 (1977).
36. Weedman, D. W. *A. Rev. Astr. Astrophys.* 15, 69–95 (1977).
37. Khachikian, E. Y. & Weedman, D. W. *Astrophys. J.* 192, 581–589 (1974).
38. Fairall, A. P. *Mon. Not. R. astr. Soc.* 180, 381–400 (1977).

Structure of taenite in two iron meteorites

SUPERSTRUCTURE in taenite of the iron meteorite, Cape York, has already been reported^{1,2}. We have now studied taenite (f.c.c. iron–nickel alloy) from another iron meteorite, Toluca, using Mössbauer and X-ray techniques. Here we report the main results of these investigations and compare the data with those reported previously^{1,2}.

The Mössbauer spectra were recorded on a conventional spectrometer, the source was 10 mCi ⁵⁷Co in palladium. The taenite lamellae used as absorbers were obtained from the Geological Museum of the University of Copenhagen. These lamellae were prepared from the iron meteorite Toluca by Cohen about 80 yr ago. The preparation technique was probably by treatment of bulk material with diluted HCl (ref. 3). The thickness of the lamellae is of the order of magnitude 50 µm and typical areas are a few square millimeters. Each Mössbauer absorber contained several lamellae. A typical Mössbauer spectrum is shown in Fig. 1. The spectrum, recorded at room temperature (18 °C), is a superposition of two spectra: (1) a central line from a paramagnetic γ-phase; (2) an asymmetric six-line spectrum from the ordered phase, that means FeNi with the L10 structure^{1,2} (see Fig. 2). The asymmetry in the spectrum is due to a quadrupole splitting, caused by the non-cubic surroundings of the iron atoms. As in Cape York we observe narrow lines, indicating that all iron nuclei have identical surroundings. The magnetic hyperfine field H_i and the electric quadrupole shift ϵ are the same as found in Cape York^{1,2} ($H_i = 288$ kG, $\epsilon = 0.20$ mm s⁻¹).

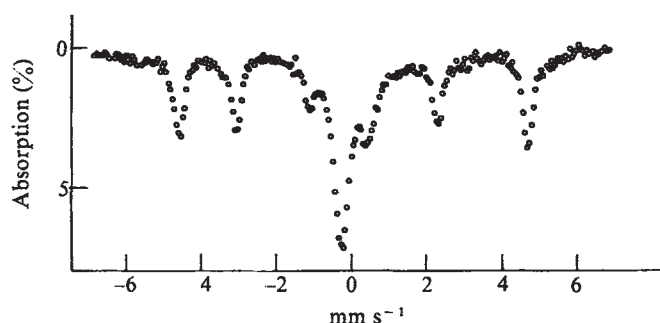


Fig. 1 Mössbauer spectrum of taenite from the iron meteorite Toluca. The spectrum was recorded at room temperature (18 °C). (Source, ⁵⁷Co in Pd).

A French group working on an ordered phase FeNi in the iron–nickel system^{4–7} found that the values of the magnetic hyperfine field H_i and the electric quadrupole shift ϵ are $H_i = 288$ kG and $\epsilon = 0.20$ mm s⁻¹. We conclude, therefore, that taenite from the meteorites Toluca and Cape York contain the ordered FeNi phase.

We have also studied the taenite lamellae from Toluca by X-ray techniques. Oscillating crystal photographs show that the lamellae are essentially single crystals and that they all

contain the superstructure L10 of FeNi. For two of the lamellae, Weissenberg photographs were obtained by oscillating the sample around a [110] direction, the oscillation interval being $\pm 90^\circ$ around a position where the incident X-ray beam is in the [111] direction, which is perpendicular to the plane of the lamella. From the zero layer and first-layer line Weissenberg photographs we have constructed the reciprocal lattice of the superstructure (see Fig. 3). The reciprocal lattice corresponds to the L10 structure with all three stacking directions about equally represented. If the ordered phase were absent, the first layer line Weissenberg photograph should show no reflections at all. However, the photograph shows 32 reflections which can all be indexed as superstructure reflections from the L10 structure. In the first as well as in the zero layer line exposure none of the expected reflections are missing.

From an X-ray exposure of two lamellae using a 114 mm camera with the film in Straumanis position we can conclude that the c/a ratio for the face centred tetragonal lattice of the ordered phase is very close to 1 ($|c-a/a| < 5 \times 10^{-4}$). It is also interesting that the lattice parameters of the superstructure and the disordered γ-phase in the lamellae are very nearly equal, both being $a = 3.582 \text{ Å} \pm 0.002 \text{ Å}$.

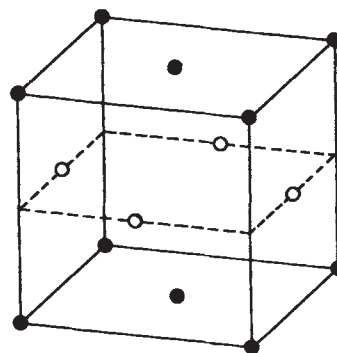


Fig. 2 The unit cell of the superstructure FeNi (L10). ●, Fe; ○, Ni.

Apart from FeNi the compound FeNi₃ is the only known ordered structure in the iron–nickel alloy system. This compound, having the L12 structure, would also give the reflections we observe. However, the lattice constant of FeNi₃, $a = 3.552 \text{ Å}$, is incompatible with the value given above and also the Mössbauer spectrum of FeNi₃ is clearly different showing a quadrupole splitting very close to zero ($\epsilon \approx 0$ mm s⁻¹) as expected for a cubic crystal^{8,9}.

A comparison of the measured value of the lattice parameter, $a = 3.582 \text{ Å}$, with published unit cell edges of disordered γ iron–nickel alloys¹⁰, indicates that the disordered γ-phase in the lamellae contains less than 28% Ni. This agrees with our Mössbauer results: from Fig. 1 it is seen that the disordered γ-phase is paramagnetic and it is known that this property is found in γ iron–nickel alloys only when the Ni content is lower than about 30% (ref. 11).

The X-ray investigations of taenite from Cape York have been supplemented with Weissenberg exposures. These show the same superstructure reflections and, furthermore, the lattice constants are also $c \approx a = 3.582 \text{ Å} \pm 0.002 \text{ Å}$.

It is remarkable that the taenite of both Toluca and Cape York have differentiated into two phases with equal lattice constants, a fact substantiated by the X-ray results. The same conclusion is reached from the Mössbauer spectra of thin lamellae^{1,2} (containing no α-phase).

This is surprising because the well known M-curve for the varying Ni concentration¹² across the taenite lamella would lead us to expect a continuously varying lattice constant and broad Mössbauer lines. The absence of both features suggests a new interpretation of the M-curve for thin taenite lamellae.

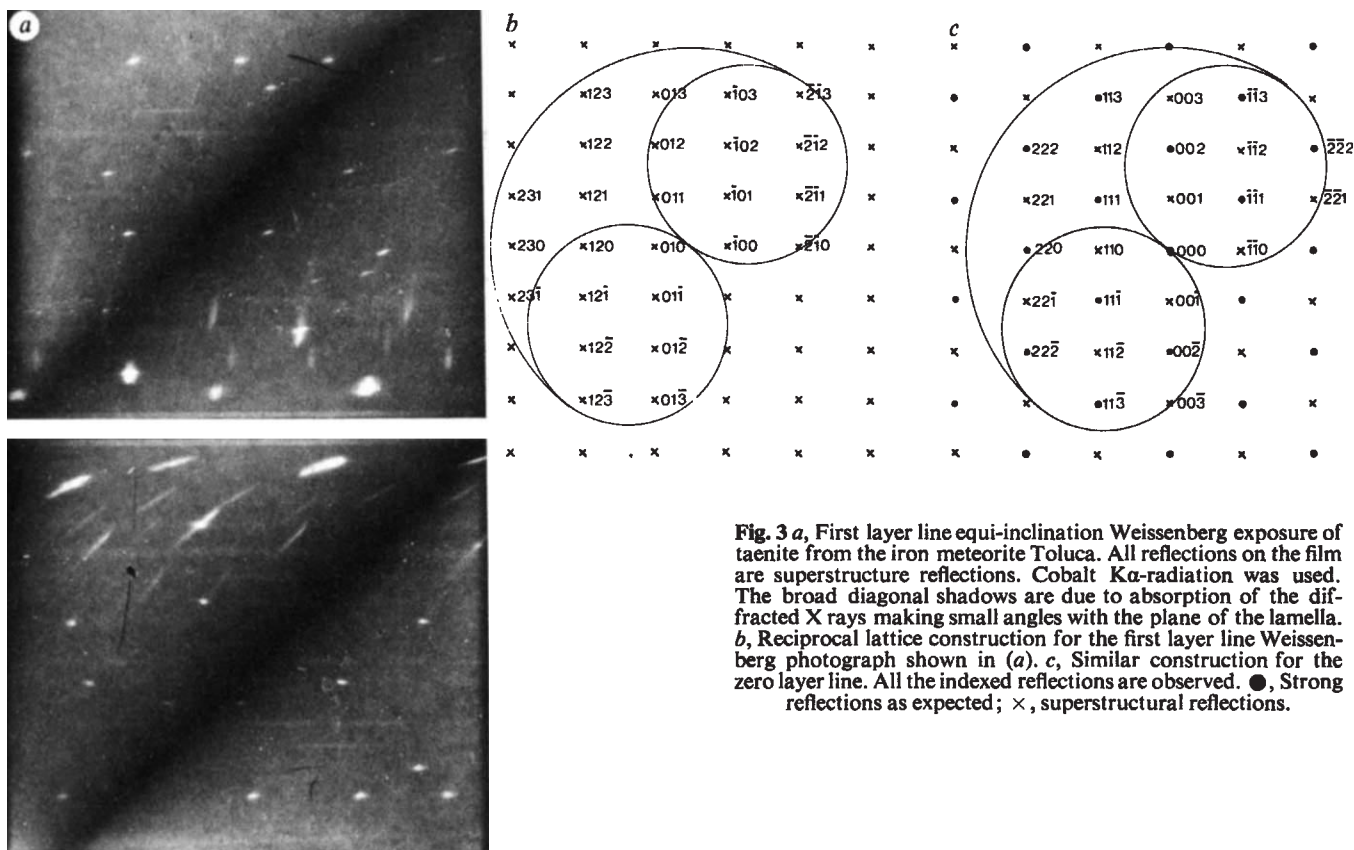


Fig. 3 *a*, First layer line equi-inclination Weissenberg exposure of taenite from the iron meteorite Toluca. All reflections on the film are superstructure reflections. Cobalt $K\alpha$ -radiation was used. The broad diagonal shadows are due to absorption of the diffracted X rays making small angles with the plane of the lamella. *b*, Reciprocal lattice construction for the first layer line Weissenberg photograph shown in (*a*). *c*, Similar construction for the zero layer line. All the indexed reflections are observed. ●, Strong reflections as expected; ×, superstructural reflections.

The varying Ni concentration is due to the occurrence of only two phases coexisting as small intergrown domains in varying amounts. The one phase, ordered FeNi, is dominant at the surface of the lamella² giving place gradually, deeper in the lamella, to an increasing amount of the disordered and paramagnetic invar alloy probably with the composition Fe (73%), Ni (27%).

This interpretation does not contradict the results obtained by microprobe analysis, as the resolution of the microprobe is $\sim 1\text{--}2\text{ }\mu\text{m}$, which is considerably larger than the dimensions expected for the domains of the two phases mentioned above.

In addition to our main investigation, we have also studied plessite from Toluca by X-ray techniques. Plessite consists of a fine mixture of taenite and kamacite. We have found that the taenite in plessite fields contains considerable amounts of the superstructure FeNi. Lin *et al.*¹³ investigated plessite from the meteorite Carlton and found that the taenite contains about 43% of Ni, indicating that this γ -phase probably also contains the ordered FeNi-phase.

Apart from the M-curve, several other characteristic features of taenite lamellae find a natural interpretation based on the model proposed above. A more detailed account, dealing also with the so-called dark etching rim, will be published later.

We thank Dr V. F. Buchwald for his interest, Dr Graff Petersen for supplying the samples investigated. J. F. A. thanks the Danish Natural SRC for support.

J. F. ALBERTSEN

Physics Laboratory, I. H. C. Oersted Institute,

G. B. JENSEN

Department of Electrophysics,
Technical University of Denmark,
DK-2800 Lyngby, Denmark

J. M. KNUDSEN

Physics Laboratory, I. H. C. Oersted Institute,
Universitetsparken 5,
DK-2100 Copenhagen O., Denmark

Received 20 February; accepted 28 March 1978.

1. Petersen, J. F., Aydin, M. & Knudsen, J. M. *Phys. Lett.* **62A**, 192-194 (1977).
2. Albertsen, J. F., Aydin, M. & Knudsen, J. M. *Physica Scripta* (in the press).
3. Cohen, E. *Meteoritenkunde* (I. Heft, Stuttgart, 1894).
4. Pauleve, J., Dautreppe, D., Laugier, J. & Neel, L. C. *r. Acad. Sci.* **254**, 965-968 (1962).
5. Cros, Y. & Pauleve, J. *J. Phys.* **31**, 459-470 (1970).
6. Neel, L., Pauleve, J., Pauthenet, R., Laugier, J. & Dautreppe, D. *J. appl. Phys.* **35**, 873-876 (1974).
7. Billard, L. & Chamberod, A. *Solid St. Commun.* **17**, 113-118 (1975).
8. Drijver, J. W., van der Woude, F. & Radelaar, S. *Phys. Rev. B* **16**, 985-992 (1977).
9. Drijver, J. W., van der Woude, F. & Radelaar, S. *Phys. Rev. B* **16**, 993-1000 (1977).
10. Ramsden, A. R. & Cameron, E. N. *Am. Mineral.* **51**, 37-55 (1966); **51**, 1544-1545 (1966).
11. Johnson, C. E., Ridout, M. S. & Cranshaw, T. E. *Proc. phys. Soc.* **81**, 1079-1090 (1963).
12. Buchwald, V. F. *Handbook of Iron Meteorites, their history, distribution, composition and structure*, 1-3 (University of California Press, 1975).
13. Lin, L. S., Goldstein, J. I. & Williams, D. B. *Geochim. cosmochim. Acta* **41**, 1861-1874 (1977).

Pressure-melting within a glacier indicated by the chemistry of re-gelation ice

IN theories explaining glacier basal sliding the transfer of water to form re-gelation ice is usually considered. However, the origin of this water has not been clearly demonstrated. In the case studied here, this water is squeezed out from the basal ice mass before re-freezing at the interface. This is indicated by the cationic composition of re-gelation ice.

Pressure-melting within the basal ice mass, as distinct from processes at the ice-bedrock interface, can be responsible for the formation of excess water in zones of high-pressure ice upstream of obstacles. If, as Robin¹ suggests, this water is squeezed out of the ice by the pressure, we have a simple heat pump, that will tend to cool the basal ice. The latent heat from the additional melting is removed with the water. Downstream of the obstacle,

basal ice temperature will be below the pressure melting point and the squeezed water can re-freeze at the interface to form re-gelation ice. The effect will be associated with large bed protuberances as well as with small ones and is completely different from the pressure-melting mechanism developed by Weertman². It is likely to occur under an alpine glacier and the question arises if it can be shown to occur.

The chemical composition in major cations of re-gelation ice at the ice-gneissic bedrock interface of the Tsijiore Nouve glacier in the Swiss Alps was determined and compared with that of other samples taken in the same environment (glacier ice, supraglacial meltwater, subglacial water and proglacial meltstream). Many precautions were taken in order to reduce the contribution from contamination to the background noise level of the atomic absorption spectrophotometer and to control chemical changes caused by the melting of the samples.³ Table 1 shows that re-gelation ice, which is bubble-poor ice containing in places thin debris layers and some rock fragments, is distinctly different from glacier ice by a higher content and a higher Na+K/Ca+Mg ratio.

Meltstream does not reach such a high Na+K/Ca+Mg ratio (Table 1). Good negative correlations exist between alkali or alkaline earth concentration and discharge in the proglacial stream just in front of the glacier (Fig. 1). The regression line for Ca+Mg is slightly steeper than for Na+K in the logarithmic coordinate system used but this slope difference is not statistically significant. This means that Na+K/Ca+Mg keeps a nearly constant value with discharge (min.: 0.05; max.: 0.09). This maximum value is not far from those obtained on supraglacial streams (0.11–0.14).

On the other hand, experiments were carried out at 0 °C in a laboratory cold room leaving in contact for various periods (1, 2, 3, 4, 5, 10, 20, 80, 160, 320 min) the fine particles (<50 µm) of the morainic deposits from the Tsijiore Nouve glacier and melted ice from the same glacier. For each selected period, 1 g of these fine particles, separated by dry sieving on nylon, was mixed with 20 ml of meltwater, shaken and filtered on a 0.45 µm acetate of cellulose filter in a Millipore plastic vacuum set. Each experiment was carried out six times. The filtered waters were analysed by AA spectrophotometry. Sediment concentration is 50,000 p.p.m. which is one or two orders of magnitude higher than in the proglacial stream (200 to 2,000 p.p.m.) and three orders of magnitude higher than in supraglacial streams (about 10 p.p.m.). Results are given in Fig. 2. They indicate that Na+K/Ca+Mg decreases in the course of time, a value as low as 0.18 being already recorded after one minute in contact.

The two facts developed above indicate that water in contact with fine particles, in the sediment concentration range 10 p.p.m.–50,000 p.p.m. which includes most situa-

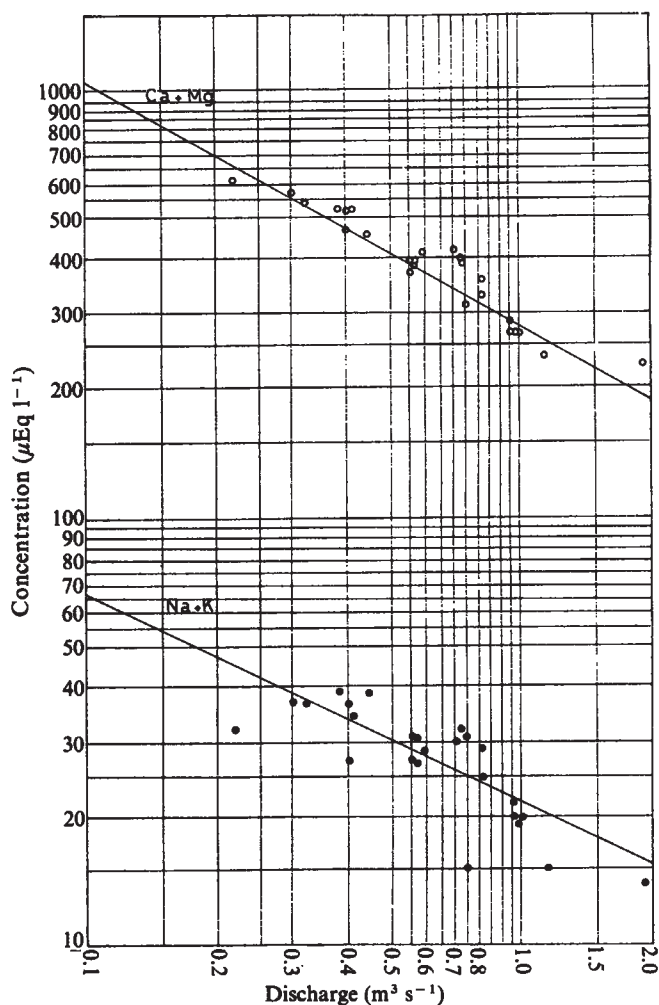


Fig. 1 Relationship between concentration and discharge in the proglacial stream of the Tsijiore Nouve glacier.

tions, does not reach; in the environment, a Na+K/Ca+Mg ratio as high as 0.58 which is the mean value for re-gelation ice. In addition selective incorporation of cations during freezing will decrease the Na+K/Ca+Mg ratio in ice as compared to water since the divalent ions are preferentially incorporated^{4,5}. Moreover a possible reduction in the Ca+Mg content of the water by subglacial carbonate deposition on the gneissic bedrock is precluded³. Melting at the interface upstream of bed protuberances and refreezing of these meltwaters downstream, as Weertman's theory implies, cannot thus explain the peculiar characteristics of re-gelation ice.

Table 1 Cationic composition of ice and water samples from the Tsijiore Nouve glacier area

	Glacier ice*			Re-gelation ice†			Meltwater‡ (supraglacial, subglacial, proglacial)	
	Mean	Min.	Max.	Mean	Min.	Max.	Min.	Max.
Na	0.8	0.3	2.0	2.9	1.3	5.2	0.3	23.1
K	0.9	0.3	1.4	2.4	0.8	3.7	0.4	16.1
Ca	4.2	1.9	9.6	6.4	3.4	10.7	4.0	536.4
Mg	2.6	1.4	7.1	3.1	1.9	4.4	0.8	74.9
Na+K+Ca+Mg	8.5	4.3	18.0	14.8	8.0	22.8	5.5	643.0
Na+K/Ca+Mg	0.28	0.09	0.49	0.58	0.34	0.85	0.05	0.14

Values are given as µEq l⁻¹.

* 18 samples.

† 13 samples.

‡ 30 samples.

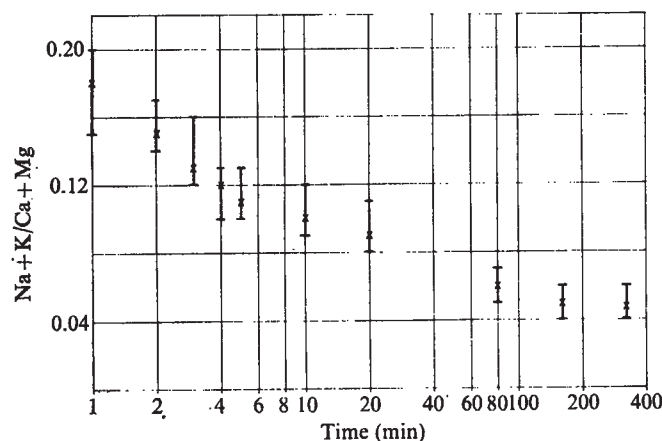


Fig. 2 Ratio of alkalis to alkaline earths as a function of time in the experiment. Time in this figure is the time elapsed between the contact of the first drop with the particles to the completed filtering.

If we now consider the mechanism discussed above, a flushing-out of ions along with the squeezed water is likely to occur. Due to a more rapid migration of alkalis which have larger diffusion coefficients than Ca or Mg, the squeezed water can be selectively enriched in Na and K. On the other hand, forcing of squeezed water across a frozen mud layer at the base can also produce an increase in Na+K/Ca+Mg ratio; indeed such a layer may be considered as an ion-exchanger membrane causing electrolyte filtration⁶. Thus the peculiar chemical characteristics of re-gelation ice are explained by refreezing of this squeezed water at the ice-bedrock interface. This gives some weight to the assertion that pressure-melting within the basal ice accompanied by meltwater squeezing and re-gelation is an active process at the base of an alpine glacier. The considerations developed above point to the need for a re-appraisal of the commonly used methods for calculating basal sliding of glaciers over surfaces of varying roughnesses.

R. SOUCHEZ
M. LEMMENS
R. LORRAIN
J.-L. TISON

Laboratoire de Géomorphologie,
Faculté des Sciences,
Université Libre de Bruxelles,
Brussels, Belgium

Received 20 February; accepted 22 March 1978.

1. Robin, G. de Q. *J. Glaciol.* **16**, 183–196 (1976).
2. Weertman, J. *J. Glaciol.* **5**, 287–303 (1964).
3. Souchez, R. & Lorrain, R. *J. Glaciol.* (in the press).
4. Malo, B. & Baker, R. *Am. chem. Soc.* **73**, 149–163 (1968).
5. Gross, G. *Am. chem. Soc.* **73**, 27–97 (1968).
6. Souchez, R. & Lorrain, R. *J. Glaciol.* **14**, 261–265 (1975).

Evidence from Iceland on geomagnetic reversal during the Wisconsinan Ice Age

THE discoveries of reversely magnetised basalts in the axial valley of the Mid-Atlantic Ridge^{1–3} and Cayman Trough⁴ have added a new facet to the debate about geomagnetic events during the recent Brunhes. Are these reversely magnetised rocks the legacy of a recent geomagnetic event or are they the product of some process of axial valley

tectonics which we do not yet understand? We have confirmed the existence of two reversely magnetised lava flows which were erupted subglacially in southwestern Iceland (Fig. 1) and seem to be direct analogues of the reversed rocks in the axial valley. We report here that a recent geomagnetic event is the only mechanism which can explain both the Icelandic and axial valley reversed outcrops.

The reversely magnetised Icelandic material, which was originally reported by Einarsson⁵ on the basis of measurement of partially orientated hand samples, forms the top of the small subglacial ridge near Maelifell, south of Thingvallavatn. It is poorly exposed in scree and outcrops in spheroidal, pillow-like structures of olivine rich basalt.

We sampled the 11 outcrops which seemed to be in place. Most inclinations steepened during cleaning in alternating fields (Fig. 2). The natural remanent intensities were unusually strong and ranged over 5–192 A m⁻¹ with an average of 50 A m⁻¹. The median destructive fields were unusually low for young subaqueous lavas; they ranged from 4 to 25 × 10³ A m⁻¹ and averaged 9.6 × 10³ A m⁻¹. Curie temperatures were run for seven specimens, and they ranged from 222 to 272 °C, with an average of 248 °C. There was no evidence of multiple Curie temperatures in any sample.

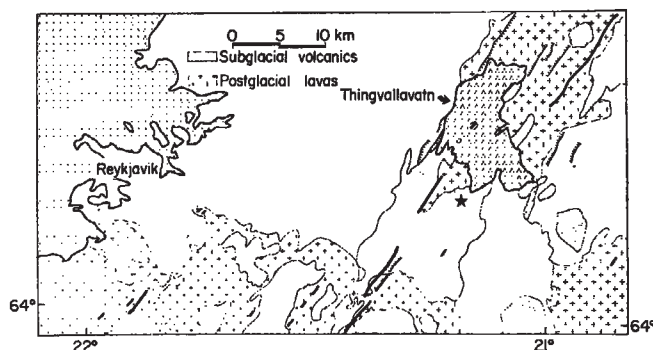
Cleaned directions, averaged by individual outcrop, are shown in Fig. 3. It is evident that there are two distinct directions defined by the data and a third direction is defined by only one outcrop. We interpret these directions as evidence that two separate lava cooling units were extruded during a time of reversed magnetic polarity. Sufficient time must have passed between the eruption of the first unit and that of the second unit to allow a large amount of secular variation to occur. The distribution of samples on the hillside indicates that unit *a* is the youngest.

Samples from pillows below the reversed units define a normal direction of *I* = +62°, *D* = 331°. The solitary shallow direction lies between the normal and reversed units and it may represent a transition direction or simply mean that the outcrop was not in place.

The reversed outcrop occurs near the top of the stratigraphic sequence of subglacial volcanics which form the Hengill complex. These volcanics are thought to have been formed under ice of Wisconsinan age⁶. Unfortunately, there are no radiometric dates to confirm this interpretation. From the nature of the subglacial pillows in the younger material of Hengill, Saemundsson (personal communication) believes that the reversed outcrop must have been erupted under a thick glacial cover, and therefore he places its age as near the end of the Wisconsinan Ice Age, probably about 20,000 yr ago. This young age is circumstantially supported by the proximity of post-glacial volcanism about 3 km north-west of the outcrop.

We have attempted to date the reversed material radiometrically using both fission tracks and potassium/argon. No fission tracks were visible in the one glass sample which

Fig. 1 Location map, southwestern Iceland. Reversed outcrop is marked by the star south of the lake. Major normal faults are indicated by hachures, and volcanic sources are indicated by heavy lines.



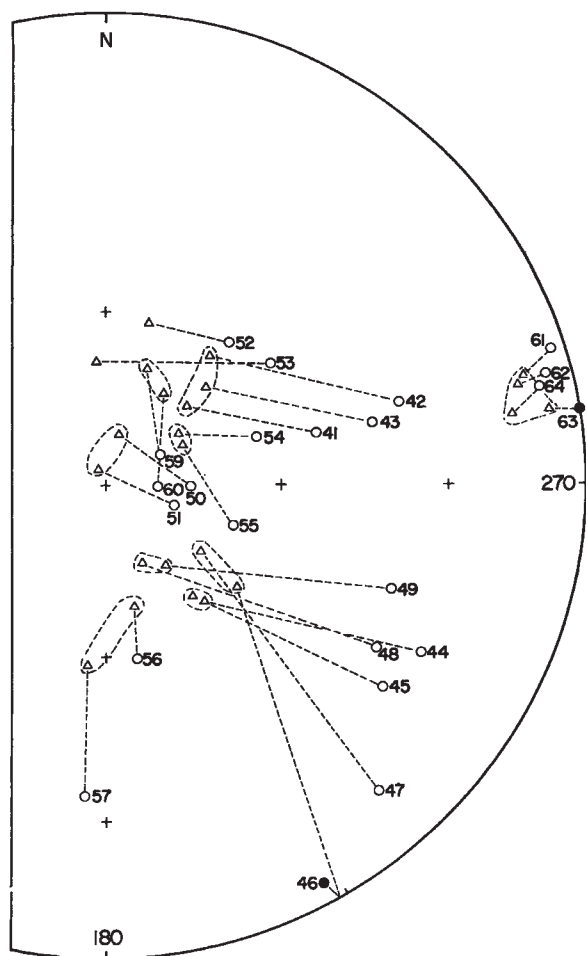


Fig. 2 Equal area plot of NRM (circles) and cleaned (triangles) directions from the individual specimens. Open symbols indicate negative inclinations. Dashed circles enclose multiple specimens from each of the 11 outcrops.

we tried to date. However, as the quality of the etching was poor, and the uranium content is very low (Gibson, personal communication), no maximum age estimate was possible. The low potassium content (0.016%) made K/Ar dating impossible. Therefore, until more suitable material for dating can be found, we can only place the age of this outcrop as upper Wisconsinan.

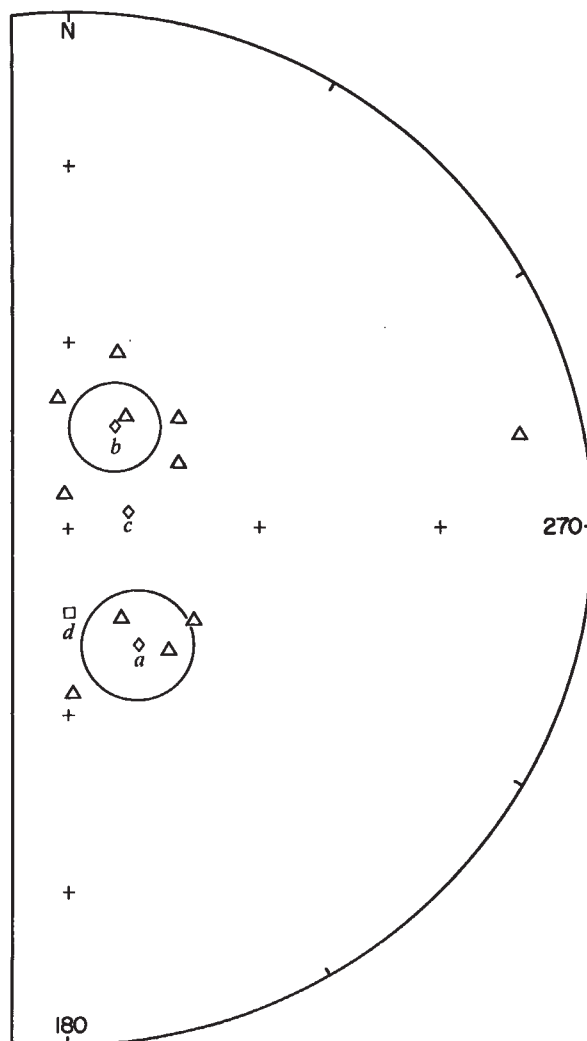
Until a firm age can be ascribed to these Icelandic reversed rocks, we suggest the name 'Maelifell event' for the geomagnetic event which we have inferred from them. On the basis of the stratigraphic relationships, the Maelifell event seems to be too young to be the Blake event⁷. It is more likely that this is the same as the Lake Mungo event^{8,9}, radiocarbon dated between 26,000 and 30,000 yr. As the Laschamp event^{10,11} has been recently redated at 33,500 yr by thermoluminescence (Valladas, personal communication), we suggest that the Laschamp, Lake Mungo, and Maelifell events are the same. Clearly, further dates are needed, especially in Iceland, to confirm this suggestion. It is not clear whether the numerous reports of geomagnetic excursions in sediment cores^{12,13} represent this event or whether they record some other geomagnetic variations.

If the geomagnetic field reversed recently, one would expect to find evidence of a large increase in atmospheric ¹⁴C due to increased production rates during the periods of low field strength associated with polarity transitions. If the transitions were each 2,000 yr long, the increase would be about 25%. Longer transitions imply larger increases. However, a recent calibration of the radiocarbon time scale

against ²³⁰Th/²³⁴U ages as old as 32,000 yr (refs 14,15) indicates no increase in ¹⁴C over 8%. If this calibration is correct, then it implies either that the Maelifell event occurred at least 40,000 yr ago, or that the transitions were much more rapid than 2,000 yr, or that there was no major worldwide reduction in geomagnetic intensity associated with the event. This last possibility suggests that the Maelifell event and other synchronous events are the products of the growth and decay of an intense non-dipole feature rather than a true dipolar reversal.

In addition to supporting the concept of a young geomagnetic event, our data have major implications for the fine-scale tectonic model of the Mid-Atlantic Ridge. Although several explanations other than a geomagnetic event have been suggested to explain the axial valley reversed occurrences², none of these is satisfactory in Iceland. The presence of this reversed outcrop in an area which is very similar tectonically to the Mid-Atlantic Ridge suggests that similar reversed units there¹⁻⁴ cooled during this reversed geomagnetic event. However, as this event has so

Fig. 3 Equal area plot of sample directions (triangles) derived from averaging the specimen directions for each outcrop. The directions fall into two groupings whose mean directions (diamonds) and 95% confidence circles are shown. The isolated direction near 270 represents one outcrop. The direction of the reversed axial dipole ($I = -76.4$, $D = 180$) is indicated by the square. The relevant statistics are: *a*, $n = 4$, $R = 3.947$, $k = 56.6$, $\alpha_{95} = 9.3^\circ$, mean $I = -68.2^\circ$, mean $D = 210.5^\circ$; *b*, $n = 6$, $R = 5.907$, $k = 53.8$, $\alpha_{95} = 7.8^\circ$, mean $I = -72.1^\circ$, mean $D = 335.2^\circ$; *c*, $n = 10$, $R = 9.914$, $k = 15.4$, $\alpha_{95} = 11.3^\circ$, mean $I = -80.5^\circ$, mean $D = 284.1^\circ$.



far only been discovered at one locality in Iceland, its duration was probably quite short. But why has it been observed so frequently on the Mid-Atlantic Ridge? A few of the reported reversed samples may be due to orientation errors, perhaps in combination with extreme secular variation, but the high frequency of occurrence cannot be discounted. Perhaps this indicates that very young lavas (<50,000 yr) form a much greater part of the surface veneer of volcanism in the axial valley than previous mapping¹⁸ has suggested. Until there is a statistically reliable estimate of the occurrences of such reversed blocks in an axial valley region, this question is debatable.

This work was supported by the NRC of Canada. We thank L. Kristjansson for his advice, the NRC of Iceland for permission to sample, B. Kristjansson, W. Clay and K. Taylor for their assistance, and H. B. S. Cooke and J. M. Hall for critically reading the manuscript.

JOHN W. PEIRCE
M. J. CLARK

Department of Geology,
Dalhousie University,
Halifax, Nova Scotia,
Canada B3H 3J5

Received 21 February; accepted 29 March 1978.

1. Ade-Hall, J. *et al.* *Can. J. Earth Sci.* **10**, 679-696 (1973).
2. Johnson, H. P. & Atwater, T. *Bull. geol. Soc. Am.* **88**, 637-647 (1977).
3. Macdonald, K. C. *Bull. geol. Soc. Am.* **88**, 541-555 (1977).
4. Clark, M. J., Hall, J. M. & Peirce, J. W. *Can. J. Earth Sci.* (submitted).
5. Einarsson, T. *Soc. Sci. Isl. Greinar* **5**, 119-159 (1976).
6. Saemundsson, K. *Acta Natur. Island.* **2**, 105 (1967).
7. Denham, C. R., Anderson, R. F. & Bacon, M. P. *Earth planet. Sci. Lett.* **35**, 384-397 (1977).
8. Barbetti, M. & McElhinny, M. *Nature* **239**, 327-330 (1972).
9. Barbetti, M. F. & McElhinny, M. W. *Proc. R. Soc. Lond. A* **281**, 515-542 (1976).
10. Bonhommet, N. & Babkine, J. C. *r. hebdo. Séanc. Acad. Sci. Paris* **264**, 92-94 (1967).
11. Bonhommet, N. & Zähringer, J. *Earth planet. Sci. Lett.* **6**, 43-46 (1969).
12. Nöel, M. & Tarling, D. H. *Nature* **253**, 705-707 (1975).
13. Vilks, G., Hall, J. M., Piper, D. J. W. *Can. J. Earth Sci.* **14**, 2007-2012 (1977).
14. Peng, T. H., Goddard, J. & Broecker, W. S. *Quat. Res.* (in the press).
15. Stuiver, M. *Nature* **273**, 271-274 (1978).
16. Ballard, R. D. & van Andel, T. J. *Bull. geol. Soc. Am.* **88**, 507-530 (1977).

Early Cretaceous basalt volcanism and initial continental rifting in Benue Trough, Nigeria

THE Benue Trough, located at the Gulf of Guinea re-entrant on the west coast of Africa (Fig. 1) is a sediment-filled 'failed rift', forming the third arm of a Cretaceous triple-rift (rrr) system that involved the South Atlantic and the Gulf of Guinea¹⁻³. The rift nature of the Benue Trough was first documented by King⁴ on geomorphological grounds, and this supposition has since been confirmed by geophysical⁵, stratigraphic⁶ and faunal evidence^{7,8}. The considerable interest shown in the stratigraphic and tectonic evolution of the Benue graben is due to its relationship to the opening of the South Atlantic and the initiation of continental

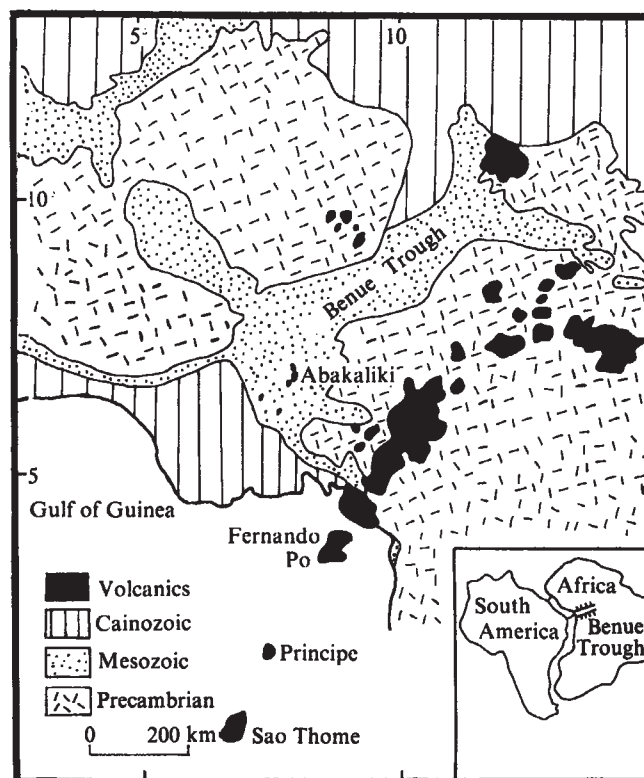


Fig. 1 Location of the Benue Trough: (inset) the Benue Trough as one arm of a Cretaceous triple junction.

drift⁹⁻¹¹. The age of the oldest Albian sediments within the trough, and their faunal biostratigraphy have been used to date the continental separation of Africa from South America⁷, and infer the history of seafloor spreading in the equatorial Atlantic region¹². New evidence is presented here on the nature of the early Cretaceous volcanism in the Benue Trough, and its petrogenesis is shown not to be due to a shortlived subduction or continental collision.

Uzuakpunwa¹³ has shown that the initiation of mid-Albian sedimentation in the Benue Trough, and by inference the opening of the South Atlantic was preceded by pre-Albian (Aptian) igneous activity involving outpouring of lavas and pyroclastic flows, probably contemporaneous with similar activity in eastern Brazil (Serra Geral lavas)²⁰. These volcanic rocks ('The Abakaliki Pyroclastics'), representing eroded remnants of more extensive flood volcanics³, are presently only well exposed at the core of anticlinal structures in the southern portion of the Benue Trough, especially around Abakaliki (Fig. 1). They are, however, older and distinct from the more widespread Upper Cretaceous to Recent alkaline rocks in the middle and upper Benue Trough. The 1,300 m of 'andesite lavas and tuffs' claimed by Burke *et al.*⁴ to be interbedded with

Table 1 Selected major and trace elements in the Abakaliki volcanic rocks

Sample no.	1	2	3	4	5	6	7	8	9	10
wt %										
SiO ₂	35.0	48.6	45.5	41.1	43.4	41.1	43.2	37.2	40.2	40.6
TiO ₂	2.70	3.15	3.00	3.73	2.73	2.31	2.18	2.35	2.83	3.39
K ₂ O	0.05	1.89	0.05	0.01	0.93	0.08	0.13	0.02	0.61	0.74
Na ₂ O	3.70	4.40	3.90	4.10	3.60	3.70	4.60	5.00	0.13	3.10
P ₂ O ₅	0.50	0.74	0.56	0.70	0.66	0.54	0.47	0.48	0.44	1.00
p.p.m.										
Zr	102	157	110	156	153	106	136	110	110	206
Y	18	18	18	20	19	21	19	18	19	22
Cr	207	219	230	313	123	384	169	193	265	211
Ni	159	130	147	200	145	258	140	143	205	132

shales at the Ikono-1 and Ogbagu-1 (oil-well) areas were found not to exist. Only a few hypabyssal sills (~10 m thick) of diabase and microdiorite belonging to the Upper Cretaceous igneous suite were encountered in the drill cores.

The occurrence of Lower Cretaceous volcanic rocks in the southern Benue Trough has been suggested^{11,13-15}, but due to the vitric nature and widespread secondary alteration (incipient spilitisation and weathering) the exact petrologic character and petrogenetic affinity of these rocks have not been clearly defined. They have been described variously as 'intermediate lavas', 'basalts' and 'andesites', and as having 'tholeiitic', 'alkaline' and 'calc-alkaline' affinities.

I have collected unweathered samples from new exposures in quarries and roadcuts in the type area, and analysed them for major and trace elements of petrogenetic significance. Petrographic examination has shown that the lavas are basaltic, commonly amygdaloidal, with intersertal to hyalopilitic texture in which microlites of albitised plagioclase are set in a fine mesocrystalline matrix of partly devitrified glass, plagioclase, chlorite, relict pyroxene (?), iron oxides, leucoxene, and minor secondary carbonate. Amygdule minerals are predominantly carbonate and chlorite. The vitric-crystal tuffs are hyaloclastic to vitroclastic, and composed of decussate laths of fine plagioclase in a matrix of spherulitic glass shards, carbonate and minor quartz.

Results of selected major and trace element analysis are presented in Table 1. The silica values range from 35 to 49% which falls within the range for basaltic rocks. The relatively low silica and variable alkalis of some of the samples can be attributed to remobilisation due to incipient spilitisation as observed in thin sections. Consequently, a standard alkali-silica plot would be of little petrogenetic significance. Instead, the TiO_2 against SiO_2 plot (Fig. 2a) was used and suggests that the rocks are alkaline. On a TiO_2 against $\text{Zr}/\text{P}_2\text{O}_5$ diagram (Fig. 2b), the alkaline affinity is further confirmed, including the typical vertical trend and virtually constant $\text{Zr}/\text{P}_2\text{O}_5$ ratios¹⁷.

As Ti, Zr and Y are immobile during alteration processes, and usually indicate different basalt types and their tectonic regimes, the method of Pearce and Cann¹⁸ has been applied to identify the petro-tectonic affinity of the Abakaliki volcanics, and the type of tectonism responsible for the formation of the Benue Trough. Using this classification (Fig. 3a), all the Abakaliki volcanics plot within the field of 'within-plate' (hotspot) basalts¹⁸. In a TiO_2 - K_2O - P_2O_5 diagram¹⁹, most of the samples lie within the field of oceanic basalts, owing partly to potash depletion. However, such dominant oceanic character in continental rocks may be attributed to volcanic activity in a zone of crustal distension and initial continental rifting¹⁹. This is consistent with the

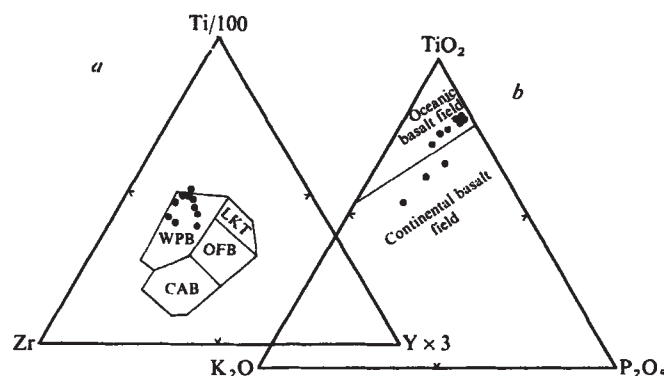


Fig. 3 a, Ti-Zr-Y plot of Abakaliki volcanic rocks. CB, 'within-plate' basalts (including hotspot basalts); OFB, ocean-floor basalts; CAB, calc-alkaline basalts; LKT, low-potash tholeiites. b, TiO_2 - K_2O - P_2O_5 plot of Abakaliki volcanic rocks.

'aulacogen' model I have proposed for the evolution of the Benue Trough¹.

This petrochemical evidence demonstrates that the early Cretaceous volcanism in the Benue Trough was not andesitic or of calc-alkaline affinity, and its petrogenesis cannot be attributed to a short-lived subduction episode^{1,2} or continental collision¹⁰. The Abakaliki basaltic rocks were related to hot-spot activity during initial continental rifting preceding the separation of Africa from South America and the influx of the sea into the Benue Trough during mid-Albian times. The alkaline rather than tholeiitic affinity of the volcanic rocks suggests that the early Cretaceous rifting was associated mainly with vertical crustal movements with little or no lateral separation as confirmed by the absence of injection dikes^{3,9}. The early Cretaceous Benue rift was probably analogous to the present-day Ethiopian rift—a system that is destined as a 'failed rift' arm or aulacogen with no possibility of undergoing crustal spreading.

I thank M. O. Oyawoye for encouragement, and S. W. Petters, W. D. Goodfellow and E. E. Ukpong for useful discussions. Sample analyses were undertaken by the Geological Survey of Canada. Fieldwork was supported by grants from the University of Ibadan.

MOSES A. OLADE*

Geology Department,
University of Ibadan,
Ibadan, Nigeria

Received 3 March; accepted 11 April 1978.

*Present address: Geological Survey of Canada, 601 Booth St., Ottawa, Ontario, Canada.

- Burke, K., Dessauvage, T. F. J. & Whiteman, A. J. *Nature phys. Sci.* **233**, 51-55 (1971).
- Burke, K. C. & Dewey, J. F. J. *Geol.* **81**, 406-433 (1973).
- Olade, M. A. *Geol. Mag.* **112**, 575-583 (1975).
- King, L. C. *Geol. Mag.* **87**, 535-539 (1950).
- Cratchley, C. R. & Jones, G. F. *Overseas Geol. Surv. Geophys. Paper*, **1** (1965).
- Carter, J. D., Barber, W. & Tait, E. A. *Geol. Surv. Nigeria Bull.* **30** (1963).
- Reyment, R. A. & Tait, E. A. *Phil. Trans. R. Soc.* **264**, 55-95 (1972).
- Petters, S. W. *Geol. Soc. Am. Bull.* **89**, 151-154 (1978).
- Wright, J. B. *Tectonophysics* **6**, 301-310 (1968).
- Nwachukwu, S. O. *Geol. Mag.* **109**, 411-419 (1972).
- Grant, N. K. *Geol. Soc. Am. Bull.* **B2**, 2295-2298 (1971).
- Gergren, W. A. & Hollister, C. D. *Soc. Econ. Paleont. Mineral. Spec. Publ.* **20**, 126-186 (1974).
- Uzuakpunwa, A. B. *Geol. Mag.* **111**, 65-70 (1974).
- Wilson, R. C. & Bain, A. D. *Geol. Surv. Nigeria Bull.* **8** (1925).
- Farrington, J. L. *Econ. Geol.* **47**, 485-508 (1952).
- MacDonald, G. A. & Katsura, T. J. *Petrol.* **82**, 82-133 (1964).
- Winchester, J. A. & Floyd, P. A. *Earth planet. Sci. Lett.* **28**, 459-469 (1975).
- Pearce, J. A. & Cann, J. R. *Earth planet. Sci. Lett.* **19**, 290-300 (1973).
- Pearce, T. H., Gormna, B. E. & Birkett, T. C. *Earth planet. Sci. Lett.* **24**, 419-426 (1975).
- Neill, W. M. *Nature phys. Sci.* **245**, 104-107 (1973).

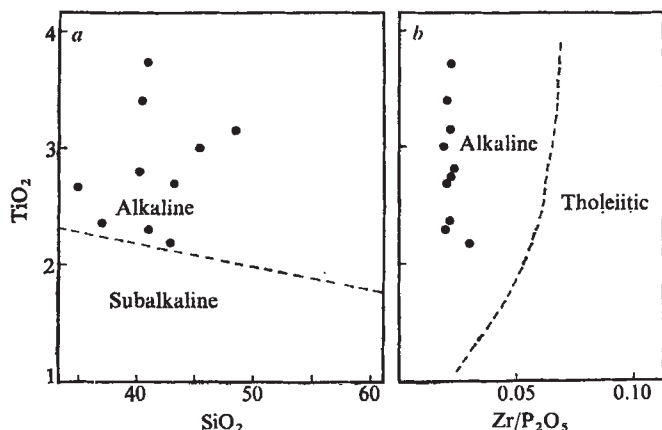


Fig. 2 Plots of (a) TiO_2 against SiO_2 ; and (b) TiO_2 against $\text{Zr}/\text{P}_2\text{O}_5$ for Abakaliki volcanic rocks.

Greenstone belts as ancient marginal basins or ensialic rift zones

EVIDENCE accumulated in several Archaean granitoid-greenstone belts indicates that greenstones were formed on, or adjacent to, pre-existing sialic crust that now consists of granitoid gneisses disrupted by younger granitoids¹⁻¹⁰. Structural studies reported here of the Eastern Goldfields Province, Yilgarn Block, Western Australia suggest a similar evolution for greenstone belts¹¹⁻¹³. We discuss the two main models which can explain the evolution of such greenstone belts. The first involves development of greenstones within ensialic rift zones^{3,14-16}, and may not have a modern analogue, whereas the second involves their development in a marginal basin environment^{17,19} similar to that developed behind modern arc systems.

A marginal basin model cannot be applied directly to Archaean terrains without considering several significant differences between greenstone belts and Phanerozoic marginal basins²⁰. Some factors, such as the lack of sheeted dyke complexes and the abundance of komatiitic lavas in greenstone belts, have been attributed to faster spreading rates and magma supply^{17,21}, but models are equivocal. For example, Tarney *et al.*¹⁷ suggest that sheeted dyke complexes are absent in the Archaean due to higher geothermal gradients, more ductile crust and reduced rifting whereas Burke *et al.*¹⁸ conclude that Archaean tectonics operated within a brittle regime to essentially present depths.

A feature not easily explained by the marginal basin model is the generation of multiple, sub-parallel greenstone belts as in the Yilgarn Block. One explanation is that arc migration led to the termination of one greenstone cycle and subsequent generation of a new basin. Such a model would necessitate a regular secular variation across the sub-parallel greenstone belts accompanied by a general secular variation across the intervening granitoid rocks as the locus of magmatism changed with arc migration. Geochronological data¹³ from the Eastern Goldfields Province of the Yilgarn Block²²⁻²⁷ do not allow direct comparison of formation ages of adjacent greenstone belts, but indicate that metamorphism and most synkinematic and postkinematic granitoid emplacement occurred over a comparatively short period between 2,700 and 2,650 Myr in widely separated areas: note the indistinguishable ages determined from 'regional' whole rock Rb/Sr isochrons in the Diemals area and the Menzies-Wiluna area (Fig. 1). This penecontemporaneous magmatism in an area of at least 210,000 km² (300 km across tectonic trends) contrasts with that of Phanerozoic marginal basins. Further, evidence from back-arc suggests a generalised sequence of emplacement from early, more basic to younger, more felsic granitoids²⁸ including late K-rich granites¹⁷ and recognisable geochemical variation towards the continent from the oceanic side²⁹. However, within a large granitoid region of the Yilgarn Block¹³ (Fig. 1) preliminary geochemical data indicate no obvious regular geochemical variation across the granitoid batholiths and field relationships negate any well-defined secular variation; the latter is extremely limited in that biotite granodiorites and adamellites predominate and tonalite/trondhjemite 'Na-rich granites' and 'K-rich granites' are rare, in contrast to previous assertions³⁰.

The granitoids so far studied from the Yilgarn Block clearly do not show the evolutionary trends expected from a back-arc system of migrating basins.

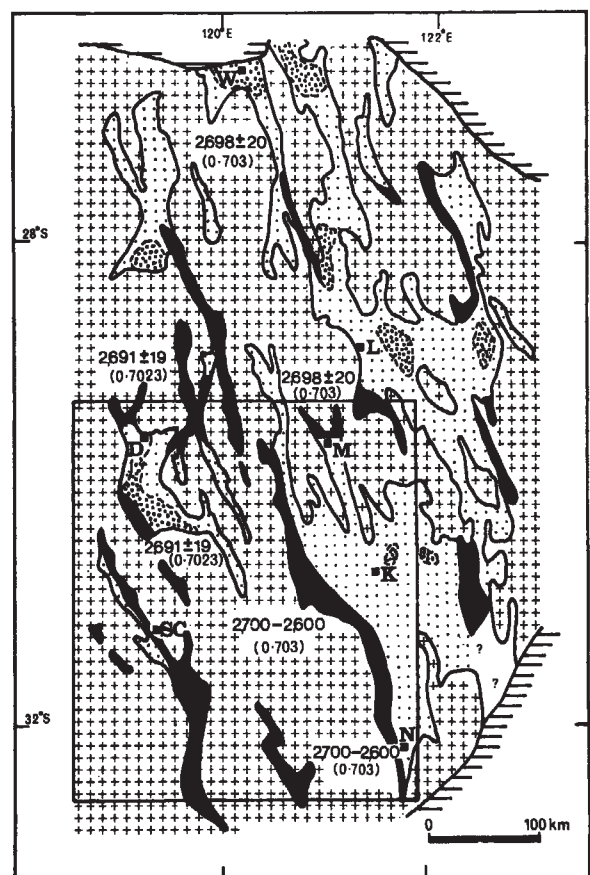
Metamorphosed felsic volcanics of the Eastern Goldfields Province also deviate considerably from those expected to develop along one margin of the greenstone belts adjacent to an arc. Instead of occurring in a linear or arcuate belt, or their deformed equivalent, isolated calc-alkaline piles are locally developed in dominantly metasedimentary sequences²⁷. More commonly restricted piles of dacitic metavolcanics are developed

without significant andesites, and may be overlain by metamorphosed ultramafic volcanic piles.

Similarly, the metamorphic pattern of the greenstone belts¹¹ does not have precise Phanerozoic equivalents. In particular, linear zones of high-grade metamorphism, intense ductile deformation and elongate, synkinematic granitoid diapirs marginal on many greenstone belts^{12,13} (Fig. 1) are not present in Phanerozoic marginal basins, although heterogeneous deformation¹⁷ is common. Although the possibility of continuous greenstone cover cannot be ruled out, these bordering linear zones have influenced our interpretation that the greenstone belts formed in ensialic rift zones whose margins were important crustal fractures that controlled both development and dynamothermal metamorphism of the greenstone belts¹³. A combination of the present limited depth of greenstone belts (< 14 km) and the preservation of very low-grade metamorphosed greenstones argue against underplating and imply a sialic floor to rift zones¹³. The observed patterns of volcanism and metamorphism have been related to the development and subsequent waning of 'hot-spots' analogous to mantle plumes¹³, and similar models have been based on the petrochemistry of Yilgarn metavolcanic rocks^{16,27}.

Constraints on an ensialic rift model are provided by Archaean deformational histories. In the Eastern Goldfields Province, deformation of the greenstones involved considerable horizontal shortening perpendicular to the long axis of the greenstone belts, as occurs in other terrains^{18,32,33} with only

Fig. 1 Map of Eastern Goldfields Province, Western Australia showing distribution and age of granitoids and distribution of metamorphic domains in greenstone belts. Numbers represent granitoid ages based on 'regional' isochrons in Myr with initial ⁸⁷Sr/⁸⁶Sr ratios shown in brackets. Area of granitoid study is outlined. +, granitoids; black area, dynamic-style, high-grade metamorphic domains with synkinematic diapirs; small dots, medium- to low-grade metamorphic domains; large dots, very low grade metamorphic domains. D, Diemals; K, Kalgoorlie; L, Leonora; M, Menzies; N, Norseman; SC, Southern Cross; W, Wiluna.



minor late deformation accompanying emplacement of synkinematic granitoid diapirs. Deformation was not related to vertical tectonics, although development of a compressional regime may have been contemporaneous with initiation of granitoid uprise¹³. Thus, the ensialic rift model, taken alone, fails to provide an adequate deformational mechanism, whereas the continental migration of an arc in a marginal basin model does. In this context a modified marginal basin model needs reconsideration.

However, if explanations for the development of parallel penecontemporaneous rift systems in a back-arc environment are sought within the context of modern plate tectonic theory there is again an unsatisfactory solution. For instance, theoretical studies of convection cells^{20,34-38} suggest that with steeper geothermal gradients and/or faster spreading rates, generally accepted for an Archaean lithosphere^{18,39}, a second order convection may occur forming horizontal rolls normal to the major convection: that is, parallel to plate motion. This may result in the penecontemporaneous development of several 'hot lines' which could explain the general volcanic cycle within greenstone belts, including the possibility of relatively shallow depths of melting for komatiites⁴⁰, the distribution and chronology of granitoids and the linear control of dynamo-metamorphism in the Yilgarn Block. However, any arc migration related to plate motion would be expected to produce maximum shortening along the lengths of the greenstone belts rather than across them as is observed. Thus, this modified marginal basin model is also inadequate to explain greenstone deformation patterns.

We have tried to stress serious objections to any modified plate tectonic model that attempts to explain the sub-parallel Yilgarn greenstone belts as a series of marginal basins, and particularly a series developed by oceanwards migration of an arc system: our data equally argue against greenstones as fossil arc systems or main ocean floors which are in any case unlikely to be preserved²⁰. All available evidence points to the development of the Yilgarn greenstone belts on sialic crust, probably in rift zones and possibly related to several penecontemporaneous hot lines, but their interpretation in terms of convection rolls developed parallel to major plate motion does not explain the observed greenstone deformation.

There seem to be insufficient constraints from detailed studies of Archaean terrains to define tectonic models of Archaean evolution and similarly insufficient data, particularly on convection models, to define comparative Archaean tectonic environments theoretically, based on modified plate tectonic models. Our study of the Yilgarn Block does, however, emphasise some important differences between modern tectonic settings and proposed tectonic environments based on modified plate models and those in which greenstone belts formed. Hence, currently proposed plate tectonic models are not adequate to explain all aspects of Archaean evolution. There is a growing bank of Archaean petrochemical data, but its use is limited unless it is integrated with field-based studies. Likewise, geochemical criteria must be used cautiously as indicators of ancient environments by direct comparisons with volcanics of modern tectonic settings. Data on Archaean granitoids and deformational histories and metamorphic patterns of greenstone belts are particularly important in constraining tectonic models.

This project was funded by the Australian Mineral Industries Research Association, the Australian Research Grants Committee and the University of Western Australia Research Committee. We thank P. J. Coleman, H. A. Doyle, J. Tarney, J. V. Watson and B. F. Windley for useful comments.

D. I. GROVES
N. J. ARCHIBALD
L. F. BETTENAY
R. A. BINNS

Geology Department,
University of Western Australia,
Nedlands, Western Australia

Received 8 February; accepted 22 March 1978.

1. Bliss, N. W. & Stidolph, P. A. *Trans. geol. Soc. S.Afr.* Sp. Publ. 2, 305-333 (1969).
2. Hunter, D. R. *Trans. geol. Soc. S.Afr.* 73, 107-150 (1970).
3. Hunter, D. R. *Precamb. Res.* 1, 259-294 (1974).
4. Stowe, C. W. *Geol. Soc. Austr. Sp. Publ.* 3, 377-383 (1971).
5. Stowe, C. W. *J. geol. Soc. Lond.* 130, 411-426 (1974).
6. Bickle, M. J., Martin, A. & Nisbet, E. G. *Earth planet. Sci. Lett.* 27, 155-162 (1975).
7. Hawksworth, C. J., Moorbath, S. & O'Nions, R. K. *Earth planet. Sci. Lett.* 25, 251-262 (1975).
8. Coward, M. P., Linken, B. C. & Wright, C. I. in *The Early History of the Earth* (ed. Windley, B. F.) 323-330 (Wiley, New York, 1976).
9. Shackleton, R. M. in *The Early History of the Earth* (ed. Windley, B. F.) 317-321 (Wiley, New York, 1976).
10. Barton, J. M., Fripp, R. E. P. & Ryan, B. *Nature* 267, 487-490 (1977).
11. Binns, R. A., Gunthorpe, R. G. & Groves, D. I. in *The Early History of the Earth* (ed. Windley, B. F.) 303-313 (Wiley, New York, 1976).
12. Archibald, N. J. & Bettenay, L. F. *Earth planet. Sci. Lett.* 33, 370-378 (1977).
13. Archibald, N. J., Bettenay, L. F., Binns, R. A., Groves, D. I. & Gunthorpe, R. J. *Precamb. Res.* 6, 103-131 (1978).
14. Windley, B. F. *Phil. Trans. R. Soc. Lond.* A273, 321-341 (1973).
15. Condie, K. C. & Hunter, D. R. *Earth planet. Sci. Lett.* 29, 389-400 (1976).
16. Nesbitt, R. W. & Sun, S. S. *Earth planet. Sci. Lett.* 31, 433-453 (1976).
17. Tarney, J., Dalziel, I. W. D. & de Wit, M. J. in *The Early History of the Earth* (ed. Windley, B. F.) 131-146 (Wiley, New York, 1976).
18. Burke, D., Dewey, F. F. & Kidd, W. S. F. in *The Early History of the Earth* (ed. Windley, B. F.) 113-129 (Wiley, New York, 1976).
19. Condie, K. C. & Harrison, N. M. *Precamb. Res.* 3, 253-271 (1976).
20. Windley, B. F. *The Evolving Continents* 24-64 (Wiley, New York, 1977).
21. Dalziel, I. W. D., de Wit, M. J. & Palmer, K. F. *Nature* 250, 291-294 (1974).
22. Turek, A. thesis, Australian National Univ. (1966).
23. Arriens, P. A. *Geol. Soc. Austr. Sp. Publ.* 3, 11-24 (1971).
24. Turek, A. & Compston, W. *Geol. Soc. Austr. Sp. Publ.* 3, 72 (1971).
25. Oversby, V. M. *Geochim. cosmochim. Acta* 39, 1107-1125 (1975).
26. Roddick, J. C., Compston, W. & Durney, D. W. *Precamb. Res.* 3, 55-78 (1976).
27. Hallberg, E. A., Johnston, C. & Bye, S. M. *Precamb. Res.* 3, 111-136 (1976).
28. Cobbing, E. J. & Pitcher, W. S. *J. geol. Soc. Lond.* 128, 421-460 (1972).
29. Coleman, P. J. *Earth Sci. Rev.* 11, 47-80 (1975).
30. Glikson, A. Y. & Lambert, I. B. *Tectonophysics* 30, 55-89 (1976).
31. Williams, I. R. *Geol. Surv. Western Australia A. Rep.* 1973 53-59 (1974).
32. Ramsay, J. G. *Trans. geol. Soc. S.Afr.* 66, 353-398 (1963).
33. Coward, M. P. & James, P. R. *Precamb. Res.* 1, 235-258 (1974).
34. Sun, S. S. & Hanson, G. N. *Geology* 3, 297-302 (1975).
35. Williams, H. R. *Nature* 266, 163-164 (1977).
36. Richter, F. M. *Rev. Geophys. Space Phys.* 11, 223-287 (1973).
37. Richter, F. M. & Parson, B. J. *J. geophys. Res.* 80, 2529-2541 (1975).
38. McKenzie, D. P. & Richter, F. M. *Sci. Am.* 235, 76-89 (1976).
39. Sinha, A. K. *Chem. Geol.* 18, 215-225 (1976).
40. McIver, J. R. *Contrib. Mineral. Petrol.* 51, 99-118 (1975).

Compensatory and conditioned feeding responses to scheduled glucose infusions in the rat

It is usually assumed that absorbed nutrients have effects which contribute to the control of food intake. Indeed, we have suggested that the rate of flow of energy to working tissues is the basic controlling factor in appetite and satiety^{1,2}. The chief absorbed energy substrate in the rat on a laboratory diet is glucose: yet glucose administration does not always produce a decrease in energy intake which compensates for all the glucose energy administered. Such negative findings may be artefacts of failure to mimic the absorption of glucose during normal satiety; sometimes, for instance, as has been pointed out^{3,4}, undetectably low doses, unphysiological concentrations and volumes, and starved rats that are not easily satiated have been used. Furthermore, it may be necessary to produce the normal context for nutrient utilisation, such as adequate secretion of insulin. For example, the reduction in food intake during continuous intravenous infusion of glucose is at most 0.6 kcal per kcal infused: coinfusion of insulin corrects this incomplete compensation⁵. We report here that the context of the initial absorption of a meal on familiar diet permits full compensation for infused glucose by a decrease in meal size. In contrast, when an unfamiliar cue is added to the diet, conditioning of appetite^{6,7} by the infusion is sufficient to overwhelm satiety, disrupting compensation and causing faster weight gain.

We first tested whether merely short-term variation in the rate of infusion throughout the day had effects which were sufficiently similar to the transient effects which normally follow a meal, as to result in completely compensatory reduction in food intake. Infusions confined to periods close to meals should in principle be even more likely to be set within appropriate conditions, and so the effects of meal-contingent infusions were tested next. Finally, infusion of

Table 1 Effects of infusing glucose on test days

Group	N	Chow intake (g)		Mean meal size (g)		Mean no. of meals		Glucose infused (equivalent in g of chow)	Energy compensation %
		Control	Test	Control	Test	Control	Test		
1	7	20.9	13.8*	—	—	—	—	7.5	94.7
2	7	21.8	20.6*	4.35	3.80*	5.60	5.67	1.9	66.0
3	6	22.5	21.6*	2.47	2.57	9.17	8.45*	3.8	23.1†
4	3	18.9	17.7	2.02	2.10	9.83	7.75	2.9	40.3†

*Significantly different from their counterparts on control days ($P < 0.05$) by paired t -test on means of individual rats (one-tailed for intakes and two-tailed for meal sizes and no. of meals).

†Significantly different from 100% compensation for the infused energy ($P < 0.05$, paired t -test).

glucose in addition to absorption of glucose could possibly have oversatiation-like aversive effects, which would be shown by the acquisition of an aversion to associated cues. Carbohydrate-rich diets, however, have been found to condition a relative preference for their flavours when the rat starts a meal, although the amount of the rich diet eaten is limited, apparently by conditioning of a relative aversion towards the end of the meal^{6,7}. To permit conditioning of aversions or preferences, an odour was added to the diet on the days of meal-contingent intravenous glucose infusion in a third group of rats. The same conditioning procedure was used in a fourth group of rats but with infusion through an intragastric cannula.

Male Wistar rats, weighing 250–350 g, were implanted with a permanent cannula either in the jugular vein or in the stomach⁸. When food intake and body weight had returned to preoperative levels, the rats were transferred to Plexiglas cages and maintained on a 12 h/12 h lighting cycle with food and water available *ad libitum*. The cannula was then connected through a watertight swivel to a syringe-drive pump. The cages were fitted with two cups mounted on continuously monitored strain gauges. A photocell and beam across the top of each food cup detected feeding and operated a relay for the control of the infusion pump. At noon each day the food cups were weighed and refilled and the rats were weighed.

During infusion Extralabo M25 powdered food was used, which yields 3.2 kcal g⁻¹ when corrected for a faecal loss. When the food was odourised, either citral (2 ml kg⁻¹) or eucalyptol (1 ml kg⁻¹) was added. These concentrations are about equally preferred in this strain of rat⁹.

The rats in group 1 (Table 1) received intravenous infusion of 30% glucose solution or isotonic saline. A volume of 20 ml was infused daily—3 min with infusion at a rate of 1.67 ml h⁻¹ followed by 3 min without infusion. The decrease in daily food intake on days of intermittent glucose infusion was almost exactly equivalent in energy to the glucose infused. Continuous glucose infusion in this dose range produces at best 60% compensation¹. There were indications in the individual results that the compensation was better at the start of the experiment. For this reason, the experiment was repeated with a further set of rats in which the first glucose infusion day was delayed for 5 or 7 d and the sequence of glucose infusion and saline infusion days was continued for 3 weeks. Figure 1 shows the pooled results from both sets of rats. The degree of compensation decreased over time to about 50%. It is not possible to say whether this represents habituation to the glucose infusion, a transfer from internal to external control of intake as the test food becomes less novel, or some other phenomenon.

Rats in group 2 were maintained on a continuous intravenous infusion of isotonic saline (1 ml h⁻¹). When a rat began to feed and broke the light beam across the top of one of the food cups, it received a supplementary infusion at a rate of 3 ml h⁻¹ which was isotonic saline or 30% glucose solution on alternate days. The supplementary infusion was led into the main infusion system via a T-piece

at a distance above the rat such that the first glucose arrived in the vein 5 min after the onset of feeding, by which time the blood glucose concentration had begun to rise¹⁰. The compensatory decrease in daily food intake was reliable (Table 1, group 2), even though the daily total amount of glucose was much smaller than in group 1. The mean 66% compensation did not differ reliably from 100%. The intake reduction was achieved by a reliable decrease in meal size on days of meal-contingent glucose infusion without any countervailing increase in the number of meals (Table 1).

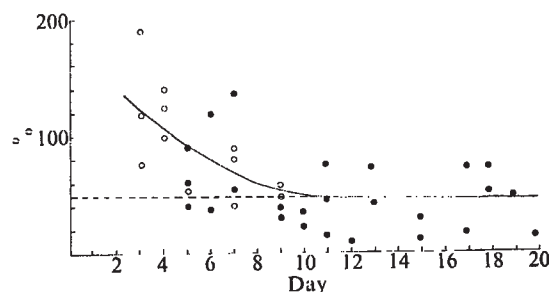


Fig. 1 Percentage compensation of food intake for glucose infused on the test day (compared with the previous day's intake) as a function of the number of days' exposure to the powdered Extralabo M25 food. ○, First test day for a rat; ●, subsequent test days.

Rats in groups 3 and 4 were provided with an olfactory cue by which they could differentiate meals during which glucose was infused from meals with control infusion. Three rats in group 3 had citral on glucose days and eucalyptol on saline days; the other three had the converse pairing. Chow having a given odour was always present in the same food cup in the same position. Although, except for these cues, the procedure was as for group 2, the result was different (group 3, Table 1). Meal size did not change and the number of meals decreased reliably. Although there was a reliable compensatory decrease in food intake, the percentage compensation was small and reliably less than 100%. Rats which survived 20 d of the schedule in good health were presented both cups of odourised food at once, each in its accustomed position on the first day of a preference test, and in reversed positions on the second day. The mean intake of the glucose-paired odour over the 2-d period was 14.7 g. Intake of the saline-paired odour was 5.5 g ($P < 0.05$).

When glucose was infused into the stomach immediately feeding started, there would naturally have been a delay before absorption occurred. Compensation in this group was poor (Table 1, Group 4) and a preference for the glucose-paired odour may have been conditioned (11.6 g compared with control odour 9.0 g).

All groups of rats showed normal slight growth, indistinguishable from that of unoperated controls, except for the rats in Group 3 which showed a significantly greater increase in body weight than controls (t -test, $P < 0.01$). This increase was reliably concentrated in glucose infusion days ($P < 0.05$).

It seems therefore that the intravenous administration of glucose, either specifically when the nutrients from a normal meal are being absorbed or intermittently perhaps mimicking absorptive transients, can change the reactions of a rat to its familiar diet such that a change occurs in its satiating power which is sufficient to provide substantial compensation. In contrast, when a new and distinctive cue is provided to be conditioned by the enrichment of the earliest stages of absorption, then the food becomes so attractive that no signs of greater satiation are seen in meal size, although some compensation is provided by a reduction in meal frequency. The conditioning effects of changes in energy supply may be more influential than any unconditioned or highly conditioned motivational effects of familiar diet. When a novelty in the diet is associated with the augmentation of the earliest stages of absorption, the new diet becomes more appetising. It does not also become more satisfying, as it does when a prolonged period of absorption follows the ingestion of a whole meal of a rich diet^{6,7}.

Thus the learning mechanisms presumed to be capable of aiding regulation⁷ can be induced artificially to operate maladaptively. This point may be relevant to the aetiology and treatment of obesity^{11,12}. It is also important to note, in current controversies over physiological satiety signals, that the conditioned preference seen here indicates that the compensatory decrease in intake seen with a familiar diet is not attributable to any aversive effect of meal-contingent infusion of glucose into the jugular vein at a rate of 15 mg min⁻¹.

P.M. was supported by an SRC CASE studentship with Unilever and a French Government CNRS grant.

Department of Psychology,
University of Birmingham,

P. MATHER

S. NICOLAIDIS
Laboratoire de Neurobiologie des Régulations,
Collège de France,
11 Place Marcelin Berthelot,
F-75231 Paris Cédex 05, France

Department of Psychology,
University of Birmingham,
PO Box 363,
Birmingham, UK

D. A. BOOTH

Received 10 January; accepted 22 March 1978.

- Booth, D. A. *Physiol. Behav.* **9**, 199–202 (1972).
- Nicolaïdis, S. *Proc. XXVI Int. Congr. Physiol. Sci. New Delhi*, **10** (1974).
- Booth, D. A. *J. comp. Physiol. Psychol.* **78**, 412–432 (1972).
- Booth, D. A. & Jarman, S. P. *J. Physiol., Lond.* **259**, 501–522 (1976).
- Nicolaïdis, S. & Rowland, N. *Am. J. Physiol.* **231**, 661–668 (1976).
- Booth, D. A. *J. comp. Physiol. Psychol.* **81**, 457–571 (1972).
- Booth, D. A. in *Food Intake and Chemical Senses* (eds Katsuki, Y. et al.) (University Park Press, Baltimore, 1977).
- Nicolaïdis, S., Rowland, N., Meile, M.-J., Marfaing-Jallat, P. & Pesez, A. *Pharmac. Biochem. Behav.* **2**, 131–136 (1974).
- Le Magnen, J. *C. r. Soc. Biol.* **150**, 136–139 (1956).
- Strubbe, J. H. & Steffens, A. B. *Physiol. Behav.* **19**, 303–308 (1977).
- Booth, D. A. in *Recent Advances in Obesity Research II* (ed. Bray, G. A.) (Newman, London, 1978).
- Booth, D. A. *Psychiat. Clin. N. Am.* (in the press).

Seroepidemiologic assessment of feline leukaemia virus infection risk for man

FELINE leukaemia virus (FeLV) is aetiological involved in naturally occurring lymphomas and leukaemias of cats. The transmission of this oncogenic virus as an infectious agent within domestic cat populations has been well-established (see ref. 1 for review). Recent studies have emphasised that, in addition to leukaemic animals, apparently healthy FeLV-infected cats may excrete high titres of infectious virus². The close environmental association of cats and man, and the known infectivity of FeLV for human cells in tissue culture^{3,4}, have raised obvious questions concerning the likelihood of horizontal transmission of this virus to humans, and the possibility that FeLV may play a part in

human disease. The availability of highly sensitive and specific radioimmunological techniques for the detection of antibodies to, and antigens of, FeLV has made it possible to systematically assess human populations, including those which might be considered to be most at risk, for immunological evidence of exposure to this oncogenic agent. In the study reported here we found no evidence of FeLV infection in man.

Radioimmunoprecipitation assays for antibodies to the major internal protein (p30) and envelope glycoprotein (gp70) of FeLV, as well as competition immunoassays for FeLV structural proteins have been described in detail^{5,6}. Sera were obtained from free-roaming, apparently healthy, domestic cats and from animals with leukaemia or lymphoma (Fairfax County, Virginia, Animal Shelter and J. Riggs, California Department of Public Health). Human sera were obtained from healthy volunteers, subjects with occupational or home exposure to FeLV-infected cats, laboratory workers with occupational exposure to RNA tumour viruses including FeLV, and patients with various neoplastic and other disorders.

To determine the prevalence of antibodies to FeLV proteins in apparently healthy cats, sera were tested for immunoprecipitation of radiolabelled FeLV p30 and gp70. Sera from 5.9% of 169 random cats precipitated quantities of either probe exceeding the background obtained with nonimmune sera. Precipitation of p30 at the highest serum concentration tested was 15% to 70% (mean 48%) and of gp70 15% to 45% (mean 25%), with titres for 10% precipitation ranging from 1:40 to 1:500. In each case, the specificity of immunoprecipitation reactions was demonstrated in competition immunoassays as illustrated in Fig. 1. Precipitation of radiolabelled FeLV p30 or gp70 was efficiently inhibited by preincubation of positive sera with detergent-disrupted FeLV grown in either feline or human cells, but was not affected by preincubation with other mammalian type-C, type-B, or type-D retroviruses. Nine antibody-negative animals were found to be antigenaemic with levels of p30 ranging from 100 to 1,600 ng ml⁻¹. That the same sera failed to react in competition radioimmunoassays for analogous proteins of unrelated viruses established the specificity of the FeLV p30 reactivities. Thus, 11% (19/169) of clinically normal cats demonstrated immunological evidence of previous or persisting FeLV infection. In contrast, sera from 72% of 80 cats with leukaemia or lymphoma contained detectable FeLV antigens (57 animals) or antibodies (1 animal). This very high prevalence of FeLV infection in cats is similar to that previously found by direct virus isolation or immunofluorescence techniques^{5,7–10}.

To determine the prevalence of antibodies to FeLV proteins in various human populations, sera from 2,074 individuals were screened in radioimmunoprecipitation assays for both p30 and gp70. Six subpopulations were identified in the study sample. Four hundred and eight healthy volunteers without known exposure to FeLV-infected animals were considered normal. The second subpopulation of 291 individuals had documented contact with FeLV-infected cats ranging from home maintenance of leukaemic animals to occupational exposure of veterinarians and animal handlers. Potential laboratory exposure to retroviruses, including FeLV, was a factor in 543 healthy adults. The fourth and fifth subpopulations, comprising 623 patients with malignancies, included 490 individuals with acute leukaemia or lymphoma and 133 patients with solid tumours. Finally, 209 patients had other, non-neoplastic disorders. Among the total population studied, no individual had detectable antibodies against either FeLV protein. Sera from all patients with haematologic malignancies and from subjects with known or potential home and occupational exposure to FeLV or FeLV-infected animals were further evaluated in competition radioimmunoassays for FeLV p30 antigen. Among this sample of 1,324 individuals considered

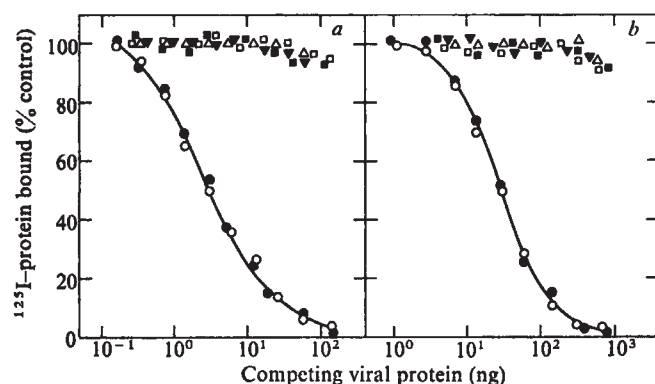


Fig. 1 Specificity of FeLV p30 and gp70 radioimmunoprecipitation by sera obtained from FeLV infected cats. Viral proteins were purified as previously described⁸ and radioiodinated by the Chloramine T method¹⁶. Antibody titres were determined by incubating serial twofold sera dilutions with 10,000 c.p.m. of ¹²⁵I-FeLV p30 or ¹²⁵I-FeLV gp70 in 0.2 ml reaction mixtures containing 10 mM Tris-HCl, pH 7.8, 0.4% Triton X-100, 1 mM EDTA, 0.5% bovine serum albumin, 0.05% sodium azide, and 10 mM (p30) or 300 mM (gp70) NaCl. After 18 h at 4 °C, anti-feline IgG was added to the reaction mixture and incubation continued an additional 4 h. Antigen-antibody complexes were pelleted by centrifugation, and the probe remaining bound in the complex quantitated in a gamma analytical counter. The specificity of sera precipitating quantities of either probe exceeding background was evaluated in competitive inhibition assays as illustrated for a representative cat serum (titres for 10% precipitation of p30 and gp70 were 1:320 and 1:160 respectively). Aliquots of 50 µl of appropriately diluted serum were pre-incubated with increasing quantities of detergent-disrupted FeLV grown in either cat (○) or human (●) cells, RD-114 (□), mouse mammary tumour virus (△), avian myeloblastosis virus (■), or Mason-Pfizer monkey virus (▼). After 6 h at 4 °C, 10,000 c.p.m. of either ¹²⁵I-FeLV p30 (a) or ¹²⁵I-FeLV gp70 (b) were added to the reaction mixture. After incubation for an additional 18 h, antigen-antibody complexes were precipitated and the radioactivity remaining bound in the pellet quantitated as described above.

most at risk of infection, none demonstrated detectable FeLV p30 antigenaemia (<50 ng ml⁻¹).

This report represents the most comprehensive sero-epidemiologic survey performed to date in efforts to assess the likelihood of FeLV transmission to humans. Despite the substantial reservoir of FeLV infected cats in the environment, no immunological evidence of FeLV infection was found in a large sample of humans potentially exposed to the virus or with malignancies pathologically similar to virus-associated feline cancers. Previous investigations have also failed to detect evidence of FeLV infection in man^{8,11-13}, although these studies were limited to relatively small numbers of selected individuals, and did not include a significant number of patients with acute leukaemia or lymphoma.

The failure to detect evidence of FeLV infection in man is striking in light of the susceptibility of cultured human cells to FeLV infection^{3,4} and in view of evidence that cats may excrete high titres of infectious virus². Humans are known to respond to immunisation with inactivated type-C viruses by producing antibodies to viral proteins¹⁴, as are subhuman primates experimentally infected with FeLV and other retroviruses¹⁵. Thus, the lack of detectable antibodies in humans to FeLV p30 and gp70 is not likely to be explained by an inability to mount an immunological response to these antigens. It is possible that FeLV may be transmitted to individuals with unusual exposure or altered host defences. Thus, continued screening of such persons for evidence of infection by FeLV remains important. However, it must be concluded on the basis of the present findings that FeLV infection occurs rarely, if at all, in the populations studied, and, as such, is likely to be a significant cause of disease in man.

We thank the following for providing human sera: J.

Gruber (Virus Cancer Program, NCI), A. Hellman (Office of Biohazards, NCI), G. Caldwell (Center for Disease Control, USPHS), and M. Gardner (University of Southern California).

JEFFERY M. KRAKOWER

STUART A. AARONSON

Laboratory of RNA Tumor Viruses,
National Cancer Institute,
Bethesda, Maryland 20014

Received 23 December 1977; accepted 13 March 1978.

- Essex, M., Stephenson, J. R., Hardy, W. D., Jr, Cotter, S. M. & Aaronson, S. A. in *Origins of Human Cancer* (eds Hiatt, H., Watson, J. D. & Winston, J. A.) 1197-1209 (Cold Spring Harbor Laboratory, New York, 1977).
- Francis, D. P. & Essex, M. *Nature* **269**, 252-254 (1977).
- Jarrett, O., Laird, H. M. & Hay, D. *Nature* **224**, 1208-1209 (1969).
- Sarma, P. S., Huebner, R. J., Basker, J. F., Vernon, L. & Gilden, R. V. *Science* **168**, 1098-1100 (1970).
- Essex, M. *et al. Int. J. Cancer* **19**, 90-96 (1977).
- Stephenson, J. R., Essex, M., Hino, S., Hardy, W. D., Jr & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1219-1222 (1977).
- Riggs, J. L., Oshiro, L. S., Taylor, D. & Lennette, E. H. *J. natn. Cancer Inst.* **51**, 449-454 (1973).
- Gardner, M. B. *et al. Int. J. Cancer* **19**, 581-598 (1977).
- Jarrett, O., Hardy, W. D., Jr, Golder, M. C. & Hay, D. *Int. J. Cancer* (in the press).
- Hardy, W. D., Jr, Old, L. J., Hess, P. W., Essex, M. & Cotter, S. *Nature* **244**, 266-269 (1973).
- Hardy, W. D., Jr *et al. Cancer Res.* **36**, 582-588 (1976).
- Schneider, R. & Riggs, J. L. *J. A. vet. Med. Soc.* **162**, 217-219 (1973).
- Gardner, M. B. *J. natn. Cancer Inst.* **46**, 281-290 (1971).
- Hersh, E. M. *et al. J. natn. Cancer Inst.* **53**, 327-333 (1974).
- Charman, H. P. *et al. J. natn. Cancer Inst.* **55**, 1419-1424 (1975).
- Greenwood, F. C., Hunter, W. M. & Glover, J. S. *Biochem. J.* **39**, 114-123 (1963).

Extratubular migration of gonocytes in the foetal rabbit testis

THE migration of primordial germ cells (PGC) from the yolk sac to the genital ridge, where they become incorporated into the developing seminiferous tubules as gonocytes, is well documented for a wide variety of mammalian species¹⁻³. Until now no further migration of PGCs out of the tubule has been observed. Here we present light and electron microscope evidence for a second migration of gonocytes out of the tubules and into the interstitium. These observations provide additional information on the mode of movement of migrating germ cells and it is also suggested that these extratubular gonocytes may be the origin of testicular teratomas in animals or embryonal carcinomas in man⁴.

The late foetal rabbit testis (29 d) consists of seminiferous cords made up of differentiating Sertoli cells and gonocytes surrounded by a limiting basement membrane (Fig. 1a). The interstitial tissue between the seminiferous cords consists of fibroblasts and Leydig cells (Fig. 1a). At this stage of gonadal development, and up to the 1-d neonate, some intratubular gonocytes move to the outside of the seminiferous cord and lie against the surrounding basement membrane where they then extend large, blunt pseudopodia through the basement membrane (Fig. 1a, b). The gonocytes then move through the holes in the membrane into the interstitial tissue, where they eventually position themselves beside the Leydig cells (Fig. 1a). The pseudopodium of each gonocyte contains numerous ribosomes, small dense granulated vesicles and microtubules (Figs 1b and 2).

Two possible mechanisms of basement membrane penetration suggest themselves. First, the release of acid hydrolases from the pseudopodium by lysosomal exocytosis could result in membrane digestion⁵. The small, dense, granulated vesicles found in the pseudopodia may be the primary lysosomes involved in this process. But this possibility would require to be tested cytochemically for the presence in them of acid phosphatase. A second mechanism for basement membrane removal is suggested by the presence of small vacuoles at the edges of the pseudopodia-containing material with a similar electron density to the membrane material (Fig. 2). These profiles suggest that the membrane could be

removed by phagocytosis.

After the initial penetration of the basement membrane the gonocyte pseudopodium seems to enlarge the breach—and a second wave of organelles, including mitochondria and rough endoplasmic reticulum, move into the pseudopodium (Fig. 2). Finally, the remainder of the cell (the nucleus, Golgi complex and the rest of the mitochondria) flows out of the seminiferous cord and moves into the interstitium (Fig. 1a). Microtubules and microfilaments are found surrounding the nucleus (Fig. 2) but seem to retract once the direction of cell migration has been clearly established. This phenomenon has now been identified in all the 10 foetuses examined between the 29th day of gestation and the 1-d neonate. No observations were made which suggested that extratubular gonocytes penetrated the basement membrane and moved back again into the seminiferous tubule.

Interstitial testicular gonocytes have also been found in all the testes examined up to 7 d postnatally, after which time they disappear. Such gonocytes retain the ability to divide extratubularly and also exhibit some of the cytochemical characteristics of the intratubular gonocytes. For example, they are positive for nonspecific alkaline phosphatase and ATPase activity⁶.

The reasons for this second gonocyte migration, which so far has not been reported as occurring in any other species, remain unclear. One possibility is that it may be a second way of reducing gonocyte numbers in the developing seminiferous tubule at the same time as the well documented intratubular gonocyte degeneration^{7,8}, the breakdown products of which after phagocytosis by Sertoli cells may be utilised in the synthesis of substances required for spermatogonial maturation⁷. Another possible reason for this reduction in gonocyte numbers in the tubule is that it may allow room for the extensive and elaborate tight junctions to develop between the Sertoli cells which form the morphological basis of the blood–testis barrier⁹.

Fig. 1 *a*, Light micrograph of a toluidine blue-stained section of part of a 29-d foetal rabbit testis. Parts of two seminiferous cords (SC) demarcated by distinct basement membranes (BM). One gonocyte (G) has extended a large pseudopodium (↑) through the basement membrane into the interstitial tissue. An interstitial gonocyte (IG) can be seen next to Leydig cells (LC). $\times 1,400$. *b*, An electron micrograph showing the initial pseudopodium (P) of an intratubular gonocyte (G) penetrating the basement membrane (↑). Dense granulated vesicles (V) and microtubules (M). $\times 6,550$.

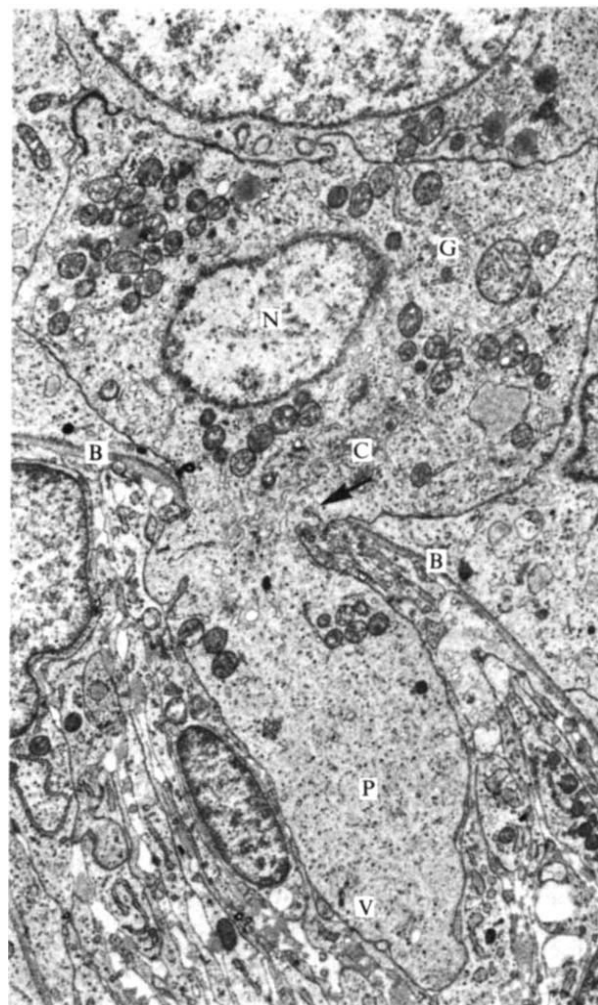
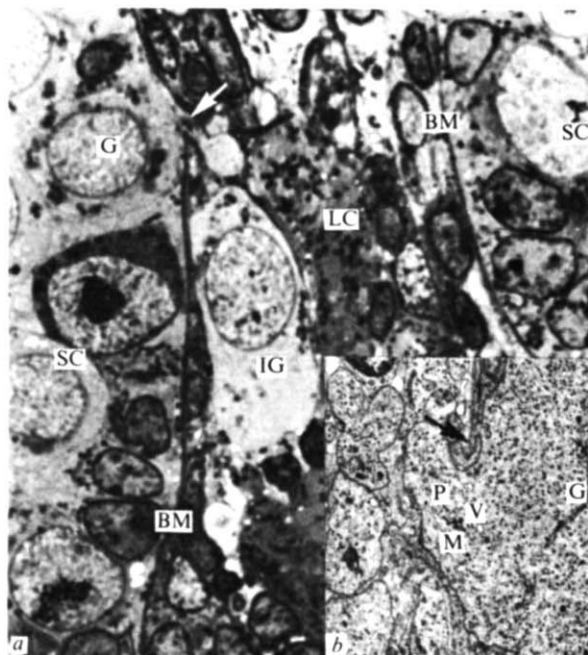


Fig. 2 An electron micrograph of part of a 29-d foetal rabbit testis. The gonocyte (G) has pushed a large pseudopodium (P) through the basement membrane (B). Dense granulated vesicles (V) can be seen in the tip of the pseudopodium which also contains free ribosomes and a few mitochondria. Apparently phagocytosed basement membrane (↑) can be seen at the margin of the pseudopodium. Microtubules and microfilaments can be seen round the nucleus (N) and Golgi complex (C). $\times 6,550$.

The mechanism of embryonic cell movement is poorly understood¹⁰ although several studies of PGC migration suggest that they move by means of pseudopodia which may be bulbous¹¹, lobate¹² and/or filiform in shape^{3,10}. The present observations support the amoeboid theory of their movement and, in particular, that their movement *in vivo* is by means of large, blunt pseudopodia and not by the filiform-shaped processes which PGCs develop when cultured *in vitro*¹⁰.

In conclusion, we suggest that even if extratubular migration is only a rare phenomenon, particularly in other species, it may nevertheless be of some pathological importance. It is widely held¹² that testicular teratomas originate in the germ cells, and this view has received support from the elegant germinal ridge transplantation experiments of Stevens¹³, who showed that mouse PGCs could develop into teratomas. The possibility thus exists that human embryonal carcinomas⁴ could also arise from extratubular gonocytes either within the testis or, if they have migrated elsewhere during foetal development in certain extragonadal sites^{12,14}.

R. P. GOULD
F. HADDAD

Department of Anatomy,
Middlesex Hospital Medical School,
London W1, UK

Received 30 November 1977; accepted 22 March 1978.

1. Chrétien, F. C. J. *Embryol. exp. Morph.* **16**, 591-607 (1966).
2. Peters, H. *Phil. Trans. R. Soc. Lond.* **B259**, 91-101 (1970).
3. Fujimoto, T., Miyayama, Y. & Fuyuta M. *Anat. Rec.* **188**, 315-330 (1977).
4. Skakkebaek, N. E. *Lancet* **i**, 516-517 (1972).
5. Witschi, E. *Contrib. Embryol. Carnegie Inst.* **32**, 67-80 (1948).
6. Haddad, F. & Gould, R. P. *Acta anat.* (in the press).
7. Gondos, B. & Hobel, C. J. *Z. Zellforsch.* **119**, 1-20 (1971).
8. Gondos, B. & Conner, L. A. *Am. J. Anat.* **136**, 23-42 (1973).
9. Dym, M. & Fawcett, D. W. *Biol. Reprod.* **3**, 308-326 (1970).
10. Heasman, J., Mohun, T. & Wylie, C. C. *J. Embryol. exp. Morph.* **42**, 149-161 (1977).
11. Zamboni, L. & Merchant, H. *Am. J. Anat.* **137**, 299-336 (1973).
12. Pugh, R. C. B. & Cameron, K. M. in *Pathology of the Testis* (ed. Pugh, R. C. B.) 199 (Blackwell, Oxford, 1976).
13. Stevens, L. C. *J. natn. Cancer Inst.* **38**, 549-552 (1967).
14. Pierce, G. B. & Beals, T. F. *Cancer Res.* **24**, 1553-1568 (1964).

Transfer of embryonic nuclei to fertilised mouse eggs and development of tetraploid blastocysts

THE transfer of nuclei from embryonic or differentiated cells into eggs is the most sensitive method of testing whether those nuclei change irreversibly during development. However, among the metazoa its use has been restricted to amphibians and insects^{1,2} because of problems associated with the microsurgery of the very small and delicate mammalian egg. A few attempts have been made to introduce somatic nuclei into unfertilised mouse oocytes or blastomeres with the help of virus-mediated cell fusion³⁻⁵, but the fate of the transplanted nucleus has not been followed beyond the cleavage following the operation. The transfer of nuclei from morula cells both by virus-mediated fusion and by microsurgery has been reported in the rabbit⁶, the evidence suggesting that such eggs can very occasionally develop to the morula stage. I describe here experiments involving a technique developed independently for use with mouse eggs. After transfer of nuclei from morula cells to fertilised one-cell eggs, more than 25% of the surviving eggs incorporated the donor genome and developed as tetraploids to the morula or blastocyst stage.

A nucleus can be transferred to a previously enucleated egg, thus replacing the egg's genome, to an artificially activated haploid egg, or to a fertilised diploid egg. If the transplanted nucleus 'takes', the resulting embryo is, respectively, diploid, triploid, or tetraploid. I decided to transfer nuclei to fertilised pronucleate eggs rather than to unfertilised eggs for three reasons. First, as microsurgery is the only routine technique for enucleating mouse eggs⁷, replacement of the egg nucleus requires double or triple puncture of the egg (depending on whether the recipient egg is unfertilised or fertilised) which considerably reduces the success rate (my unpublished observations). Second, if an egg carrying its own nucleus as well as a transferred nucleus developed, it would show that the operation had not affected viability. Third, tetraploidy in the mouse is not lethal until the second half of pregnancy¹⁰⁻¹², so that the

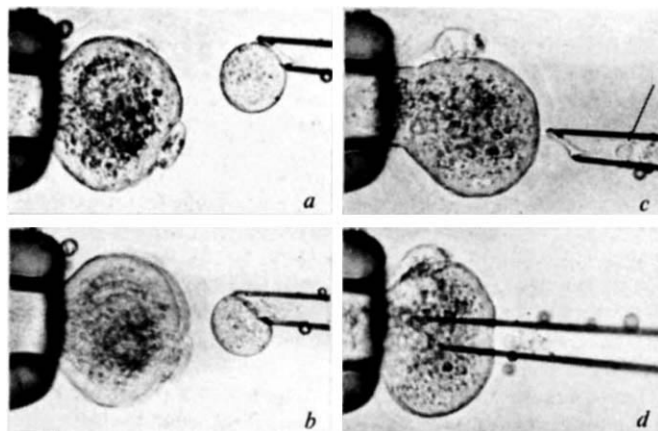


Fig. 1 The course of transfer of the morula cell nucleus into a fertilised mouse egg ($\times 360$). a, Donor cell adhered by suction to the end of the injection pipette; b, the beginning of the cell's suction into the pipette; c, the cell is sucked into the pipette and squashed and broken. Nucleus surrounded by cytoplasm lies close to the tip. The nucleolus is clearly visible (arrow); d, the pipette is inserted into the egg and the nucleus deposited in the egg cytoplasm.

activity of the transferred nucleus incorporated into a tetraploid genome can be studied during a relatively long period of development.

Fertilised one-cell eggs were obtained from 129/ter Sv female mice induced to ovulate by standard techniques and mated with A strain males. Eggs were recovered 18-20 h after injection of human chorionic gonadotrophin, and cumulus cells were removed with hyaluronidase. Donor cells originated from 12-18-cell CBA-T6T6 embryos. After the zona pellucida had been removed with Pronase, embryos were placed for approximately 20 min in phosphate-buffered saline (PBS) free of Ca^{2+} and Mg^{2+} , and then were disaggregated by pipetting. Nuclei were obtained by drawing a cell into a pipette of inside diameter smaller by one-third than that of the cell, so that the cell was broken. The tip of the pipette was shaped to facilitate insertion into eggs. The shape was obtained by breaking the tip or by treating it with hydrofluoric acid (my unpublished work). The donor nucleus and surrounding cytoplasm, in the tip of the pipette, were introduced into the egg by very slowly increasing the pressure with a Beaudouin syringe (Fig. 1a-d).

Usually the nucleus could be seen throughout the transfer operation. The whole micromanipulation procedure was basically similar to that used for enucleation of eggs and has been described in detail⁸. Most of the 166 eggs that received nuclei degenerated immediately; the rest were kept for approximately 1 h in culture medium¹³ at 37 °C, and those that looked normal were transplanted to the oviducts

Table 1 Transfer of morula cell nucleus to the mouse zygote

No. of transfers	No. of eggs which survived operation and were transplanted	Duration of development in oviduct (d)	Stage*	No. of nuclei (no. of metaphase plates)	Ploidy
166	6	4	Morula	14 (1)	2n
			Morula	12 (1)	4n
			Blastocyst	10 (1)†	4n
	9	5	Morula	10 (2)	4n
			Blastocyst	31 (1)	2n
			Blastocyst	17 (2)	4n
			Blastocyst	25 (0)	?
			Blastocyst	29 (0)	?

*Two three-cell eggs and five degenerated eggs not included.

†This figure is an underestimate due to wide scattering of nuclei on the slide.

of immature females. After 4 or 5 d the recipients were injected with colchicine (0.02 ml of 0.02% g^{-1}) and killed 2–3 h later. All 15 eggs that had been transplanted were recovered and examined in air-dried preparations¹⁴.

The eggs recovered consisted of five blastocysts, three morulae, two three-cell eggs and five degenerated eggs (Table 1, Fig. 2a). Two blastocysts and two morulae contained tetraploid metaphase plates with two T6 marker chromosomes (Fig. 2b), showing that the transferred nuclei contributed to the genome of those embryos. Two blastocysts had no metaphase plates and two embryos (a blastocyst and a morula) had diploid plates with karyotype characteristic of the recipient egg (no marker chromosomes). All embryos, including diploid ones, were retarded in development, apparently as a consequence of severe trauma caused during the injection of nuclei.

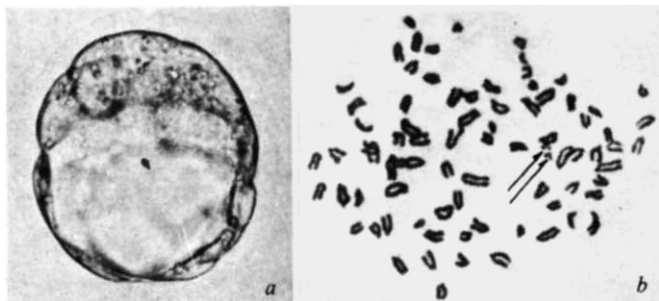


Fig. 2a, A 5-day tetraploid blastocyst ($\times 500$). b, The tetraploid metaphase plate from 13-cell morula obtained after transplantation of the morula cell nucleus. Note two T6 marker chromosomes (arrows) which originate from the transplanted nucleus.

In embryos which turned out to be diploid, the injected nucleus may have either escaped from the egg immediately after operation, been damaged and degenerated, or failed to associate with the pronuclei in contributing to the common metaphase plate of the first cleavage because of the asynchrony of recipient and donor cells. At the time of operation, recipient eggs were at the beginning of the S phase^{15,16}. The donor cells could have been in any stage of the cell cycle because they originated from 12–18-cell embryos in which cleavages are no longer synchronous. Although the exact fate of nuclei transferred in asynchrony is not known, it is likely that this factor contributes to the low success of nuclear transfer in mammals.

In my experiments only 9% of eggs survived the operation, but four of the 15 survivors developed successfully with the transferred genome. It seems, therefore, that if the egg survives the operation the chance that the nucleus will 'take' is relatively high. Although in similar experiments with the rabbit⁸ survival was much higher (72%), only 5% of eggs contained the donor nucleus and only 12 out of 428 (less than 3%) cleaved. My results show unequivocally that an embryonic nucleus can contribute its genetic material to the genome of an egg and that the resulting embryo can develop at least as far as the blastocyst stage.

This work was financed in part by the Committee of Cell Pathophysiology of the Medical Section of the Polish Academy of Sciences. I thank Professor A. K. Tarkowski and Dr A. McLaren for their comments on the manuscript.

JACEK A. MODLINSKI

Department of Embryology,
Institute of Zoology,
University of Warsaw,
Warsaw, Krakowskie Przedmiescie 26/28, Poland

Received 14 December 1977; accepted 9 March 1978.

- Briggs, R. & King, T. *Proc. natn. Acad. Sci. U.S.A.* **38**, 455–463 (1952).
- Gurdon, J. B. *The Control of Gene Expression in Animal Development* (Clarendon, Oxford, 1974).

- Graham, C. F. *Wistar Inst. Symp. Monogr.* No. 9, 19–35 (1969).
- Baranska, W. & Koprowski, H. *J. exp. Zool.* **174**, 1–4 (1970).
- Bernstein, R. M. & Mukherjee, B. B. *Nature* **238**, 457–459 (1972).
- Bernstein, R. M. & Mukherjee, B. B. *Dev. Biol.* **34**, 47–65 (1973).
- Lin, T. P. *et al. Nature* **242**, 47–49 (1973).
- Bromhall, J. D. *Nature* **258**, 719–721 (1975).
- Modlinski, J. A. *J. Embryol. exp. Morph.* **35**, 897–905 (1975).
- Snow, M. H. L. *Nature* **244**, 313–315 (1973).
- Snow, M. H. L. *J. Embryol. exp. Morph.* **34**, 707–721 (1975).
- Tarkowski, A. K. *et al. J. Embryol. exp. Morph.* **41**, 47–64 (1977).
- Whitten, W. K. *Adv. Biosci.* **6**, 129–139 (1972).
- Tarkowski, A. K. *Cytogenetics* **5**, 394–400 (1966).
- Luthardt, F. W. & Donahue, R. P. *Expl. Cell Res.* **82**, 143–151 (1973).
- Siracusa, G. & Vivarelli, J. *Reprod. Fert.* **43**, 567–569 (1975).

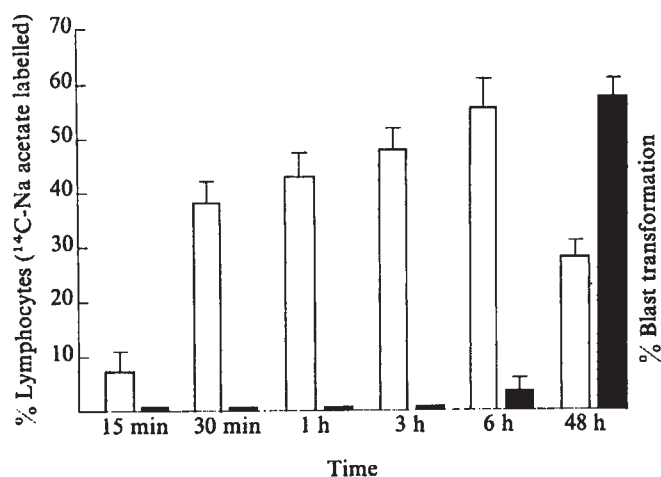
Possible use of histone acetylation in mixed leukocyte cultures to assess histocompatibility

MIXED leukocyte cultures (MLC) have been extensively used to assess histocompatibility in transplantation^{1–5}. The MLC test^{6,7} is an *in-vitro* model for determining the early recognition phase of the homograft rejection, and a correlation between the blast transformation in the MLC and the outcome of cadaver kidney transplantation has been suggested^{1–5,8}. However, the MLC test is limited in its use in cadaver kidney transplantation because of the extended period required for its completion. Therefore, a more rapid technique other than blast transformation in MLC is required to assess the potential success of kidney transplantation when cadaver donors must be used. We report here that acetylation of histones in MLC may be a time-saving technique to determine histocompatibility in cadaver kidney transplantation.

It is well known that *in vitro* stimulation of lymphocytes by mitogens such as phytohaemagglutinin (PHA) results in blast transformation as in MLC. One of the earliest events before blast transformation in PHA-stimulated small lymphocytes *in vitro* is the acetylation of histones^{9–11}. This can be detected within 15 min after stimulation of the lymphocytes with PHA and these cells eventually undergo blast transformation and mitosis. We investigated histone acetylation in MLC using leukocytes from mono- and dizygotic twins and unrelated individuals in order to determine if there is any correlation between histone acetylation, lymphoblast transformation (in MLC) and histocompatibility between donor and recipient cells.

Fresh heparinised venous blood was obtained from age and sex matched donors, and leukocytes were separated by standard procedure¹⁰. By this method a sample of blood was allowed to stand at room temperature for approximately 30–45 min and the buffy coat was resuspended in 5 ml Dulbecco's minimal essential medium (MEM) without serum. After 15 min this was slowly decanted and the supernatant reincubated; this procedure

Fig. 1 Time sequence of ¹⁴C-Na acetate incorporation and blast transformation in MLC with cells from unrelated individuals. Note that significant isotope incorporation occurred 30 min after MLC, whereas blast transformation was not appreciable until 48 h. ■, Lymphoblasts; □, labelled small lymphocytes.



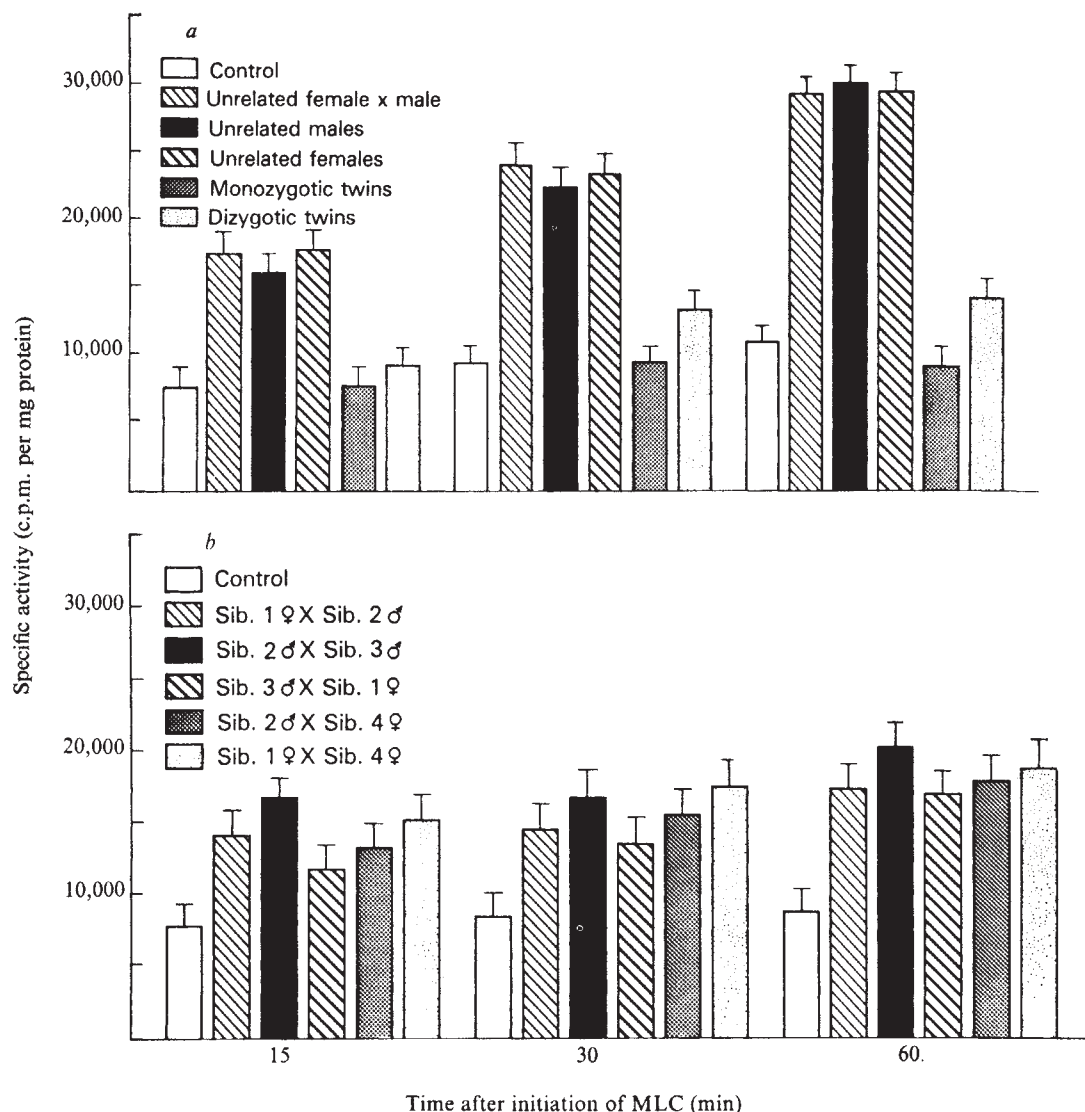


Fig. 2a. Time course of histone acetylation in MLC. Note that unrelated cultures have the highest rate of acetylation, whereas in the monozygotic twins the rate is the same as in the control cultures. These differences could be detected even within 15 min of culture. **b.** Pattern of acetylation in MLC between non-twin siblings. Note that non-twin siblings have a higher rate of acetylation compared to the mono- and/or dizygotic twins but still not as high as in the unrelated cultures.

was repeated three times. This mixture of cells was then centrifuged at 1,500 r.p.m. and the cell pellet collected. 80–85% of the polymorphonuclear neutrophils (PMN) were thus eliminated and a fairly clean mononuclear cell suspension obtained which was then used for MLC. The following combinations of MLC were set up: unrelated normal male with a normal female; unrelated normal male with a normal male; unrelated normal female with a normal female; HLA-identical monozygotic twins; dizygotic twins; the non-twin siblings. In addition, for each experiment several unmixed cultures of each individual were set up with PHA, as controls to verify the viability of the cells. Unmixed cultures without PHA were also set up to detect whether or not any endogenous stimulus, such as virus infection, existed to produce spontaneous blast transformation.

The method for labelling, isolation and purification of histones and the assay for acetylation has been described⁹. The autoradiographic method for determination of histone acetylation has been reported elsewhere^{10,11}. Blast transformation was quantitated from a series of slides containing labelled cells in autoradiographs and stained with Giemsa-9. These slides were made from cultures at various time intervals ranging from 15 min to 48 h. The slides were precoded and scored blindly by one individual.

Figure 1 shows the time course of acetylation and blast transformation in our culture system over the 48-h period. ¹⁴C-sodium acetate incorporation into the mononuclear cells as judged by autoradiography is appreciable within 15 min following initiation of MLC, as previously described for PHA-stimulated cultures¹⁰. The percentage of labelled cells reached a

peak of 6 h with a subsequent decline up to 48 h. On the other hand, blast transformation was not evident until 6 h after initiation of the MLC and it was not until 48 h that a significant number of blast cells were seen in these cultures.

The results of biochemical assay for acetylation in MLC of various combinations of donors have been presented in Fig. 2a and b. In this instance, acetylation was studied in cultures ranging from 15 to 60 min following initiation of MLC. The data indicate that there is no enhancement of detectable acetylation in cells of MLC from HLA-identical monozygotic twins, whereas unrelated cultures showed the highest acetylation (30×10^3 c.p.m. per mg protein). These counts are significantly higher than in the cultures of dizygotic twins (14×10^3 c.p.m. per mg protein). In the non-twin siblings histone acetylation occurred at a higher rate (20×10^3 c.p.m. per mg protein) than in the dizygotic twins, but was substantially lower than in the unrelated cultures.

These results suggest that histone acetylation precedes blast transformation in the MLC as in PHA-stimulated cultures and that acetylation is measurable within a very short time after initiation of the cultures (15 min). This may allow determination of histocompatibility between cadaver donor and recipient in kidney transplantation with the same degree of confidence as in MLC blast transformation. The advantage in this technique is the time saved, which is crucial to using more cadaver kidneys. Additionally, this aspect of the MLC test may make it a viable and urgently needed tool in the selection of cadaver donors.

The observed higher rate of acetylation in the non-twin siblings compared with dizygotic twins is intriguing. The

possibility of cellular exchange between the two fetuses in the same uterine environment, thus providing increased histocompatibility in the dizygotic twins, is open to investigation.

We thank Drs Joseph D. Schulman, Sheldon Orloff and Jean Deb. Butler for critical review of the manuscript, and Richard Waite and Maria Chen for technical assistance. This work was partially supported by grants from the DHEW Maternal and Child Health Service (Project no. 417), NIH (RR-05493, HD-05187 and HD-06321).

ANIL B. MUKHERJEE*

State University of New York at Buffalo,
Division of Human Genetics, 86 Hodge Avenue,
Buffalo, New York 14222

JILL D. FABRICANT†

Department of Biology,
McGill University,
Montreal, Quebec, Canada

Received 12 January; accepted 3 April 1978.

*Present address: National Institute of Child Health and Human Development, NIH Bethesda, MD 20014.

†Present address: Institute Pasteur, Paris, France.

1. Bach, J. R., Debray-Sachs, M., Crosnier, J., Kreis, H. & Dormont, J. *Clin. Exp. Immun.* 6, 821–824 (1970).
2. Opetz, G. & Terasaki, P. I. *Transplantation* 23, 375–380 (1977).
3. Hamburger, J., Crosnier, J., Descamps, B. & Rouinska, D. *Trans. Proc.* 3, 260–267 (1971).
4. McDonald, J. C., Su, F. & Hornung, M. O. *Ann. Surg.* 175, 925–932 (1972).
5. Ringden, O., Moller, E., Lundgren, G. & Growth, C. *Transplantation* 22, 9–17 (1976).
6. Bain, B., Vas, M. R. & Lowenstein, L. *Blood* 23, 108–112 (1964).
7. Bach, F. & Hirschhorn, K. *Science* 143, 813–816 (1964).
8. Morris, P. J., Kincaid-Smith, P., Ting, A., Stocker, J. W. & Marshall, V. C. *Lancet* ii, 803–804 (1968).
9. Pogo, B. G. T., Allfrey, V. G. & Mirsky, A. E. *Proc. natn. Acad. Sci. U.S.A.* 55, 805–812 (1966).
10. Mukherjee, A. B. & Cohen, M. M. *Expl. Cell Res.* 54, 257–260 (1969).
11. Mukherjee, A. B., Waite, R. G., Cohen, M. M. & Bernstein, R. *Cancer Res.* 32, 1833–1841 (1972).

Acetylcholine-induced conductance fluctuations in cultured human myotubes

THE application of acetylcholine (ACh) to chemosensitive regions of vertebrate skeletal muscle fibres elicits a conductance change in the muscle membrane which results in a net inward current flow. During a steady response to applied ACh, the ACh-evoked current fluctuates about its mean level^{1–3}. This fluctuation arises from the statistical variation in the number of open membrane channels and can be analysed to give estimates of both the conductance and the lifetime of single open channels. Analysis of drug-induced membrane current fluctuations has provided important information about the properties of cholinergic membrane channels in frog^{1–3}, chick^{4,5} and rodent^{6–8} muscle cells. We report here some properties of ACh-activated channels in human myotubes grown in tissue culture. We have also made a preliminary study of the effects of a serum from a patient with myasthenia gravis on the properties of these channels.

Human muscle cells were grown at 37 °C in either explant or dissociated cultures. The source of muscle was either 12–16-week-old fetuses or amputated adult limbs. For explant cultures, small pieces of muscle (~1 mm²) were grown on gelatin-coated coverslips in modified Eagle's medium⁹ with 10% foetal calf serum. Myotubes formed in the explant outgrowths after about 3 weeks and the cultures selected for study were aged 4–7 weeks. Dissociated muscle cultures were prepared as described for rat muscle¹⁰ and used after 8–11 d in culture. The properties of ACh channels seem to be independent of both the source of muscle tissue and the method of culture.

Experiments were made in a recording solution containing (mM) NaCl, 150; KCl, 5; CaCl₂, 10; HEPES-NaOH, 5, pH 7.4. Myotubes were voltage clamped with two bevelled microelectrodes¹¹ filled with 4 M potassium acetate (resistance 5–15 MΩ) using feedback gains of 900–1,200. The

cells studied were selected for a size and shape which would ensure approximate isopotentiality. Typical cell dimensions were 50×200 μm and cell input resistance ranged from 5 to 20 MΩ. ACh was applied iontophoretically through pipettes filled with 4 M AChCl (resistances 50–100 MΩ) by a constant current circuit isolated from the ground. The iontophoretic pipettes were positioned at some distance from the cells under study to reduce further any effects that might arise from fluctuations in the amount of ACh released. In most cases, a suitable response was elicited by either reducing or removing the steady backing current passed through the ACh pipette. The reversal potential of ACh-induced currents was estimated directly by clamping myotubes at various holding potentials and determining the potential at which no change in current was evoked by ACh. At 23 °C the reversal potential was found to be 0.6±0.6 mV (mean±s.e., six cells). Given the inherent errors in these measurements, we have taken the reversal potential to be 0 mV. Similarly a reversal potential of 0 mV was calculated for cells at 37 °C.

Application of ACh to a myotube clamped at a negative holding potential evoked an inward current associated with an increase in the variance of the current fluctuations about the mean value (Fig. 1A). Power spectra obtained from analyses of these current fluctuations were fitted reasonably well by a function of the form

$$S(f) = 2\mu V \gamma \tau / (1 + (2\pi f \tau)^2)$$

where f =frequency, τ =mean open channel time, γ =channel conductance, μ =mean ACh-evoked current and V =difference between the holding potential and the reversal potential (Fig. 1B). Such a relationship is consistent with two-state model of channel conductance proposed for ACh receptors at the frog neuromuscular junction¹. For each experiment, τ and γ were estimated by fitting a function of the above form to the estimated power spectrum by a method of least squares.

The mean channel lifetime (τ) was found to vary with the holding potential according to the equation $\tau = A \exp(-BV_1)$, where V_1 is the holding membrane potential (Fig. 2). From measurements made in 15 myotubes at 23 °C we calculate A to be 7.1±1.2 ms (mean±s.e.) and B to be 0.0090±0.0031 mV⁻¹. The value of B is similar to that previously reported for ACh receptors at frog¹ and rat² motor endplates. The mean channel lifetime at -70 mV and 23 °C was 12.2±1.3 ms (15 cells). At 37 °C and -70 mV, τ was reduced to 2.4±0.2 ms (6 cells) which is significantly different from the value at room temperature (Mann-Whitney test, $P<0.01$). If $\log \tau^{-1}$ is linearly related to the reciprocal of the absolute temperature between 23 °C and 37 °C, as is found in cultured rat myotubes (our unpublished observations), then from the Arrhenius equation we calculate that the activation energy for channel closing is 21 kcalorie M⁻¹. Our estimates of τ are considerably larger than the values reported for ACh receptors at human endplates¹² (~1.5 ms at -80 mV and 23 °C). This discrepancy probably reflects some extrajunctional-type receptor properties of ACh receptors in tissue cultured muscle cells. Other studies have shown that τ for extrajunctional receptors in denervated adult muscles is greater than that for junctional receptors^{7,13,14}.

At 23 °C the mean single channel conductance (γ) in cells clamped at -70 mV was 40.1±1.3 pS (15 cells). This estimate of γ is similar to that found in tissue cultured chick⁴ and rat myotubes (our unpublished observations), but larger than the values reported for ACh receptors at human¹², rat⁴ and frog³ neuromuscular junctions (20–25 pS) and for extrajunctional receptors in denervated frog¹³ and rat⁷ muscle (15–25 pS). Channel conductance was independent of membrane potential over the measured range of -20 to -110 mV. The mean channel conductance in cells at 37 °C was 38.8±4.7 pS (6 cells, -70 mV) which is not

significantly different from the value at room temperature (Mann-Whitney test, $P > 0.10$).

We have shown that human myotubes exposed to sera from myasthenic patients have a lowered sensitivity to ACh (ref. 15). This reduction is thought to result from some action on ACh receptors of anti-receptor antibodies found in myasthenic sera. A preliminary experiment was therefore carried out to determine if the reduced sensitivity could be attributed to altered channel properties. Myotubes were incubated with 10% myasthenic serum for 1.5 h at 37 °C, before electrophysiological examination in a serum-free recording solution at 23 °C. In such conditions, the test

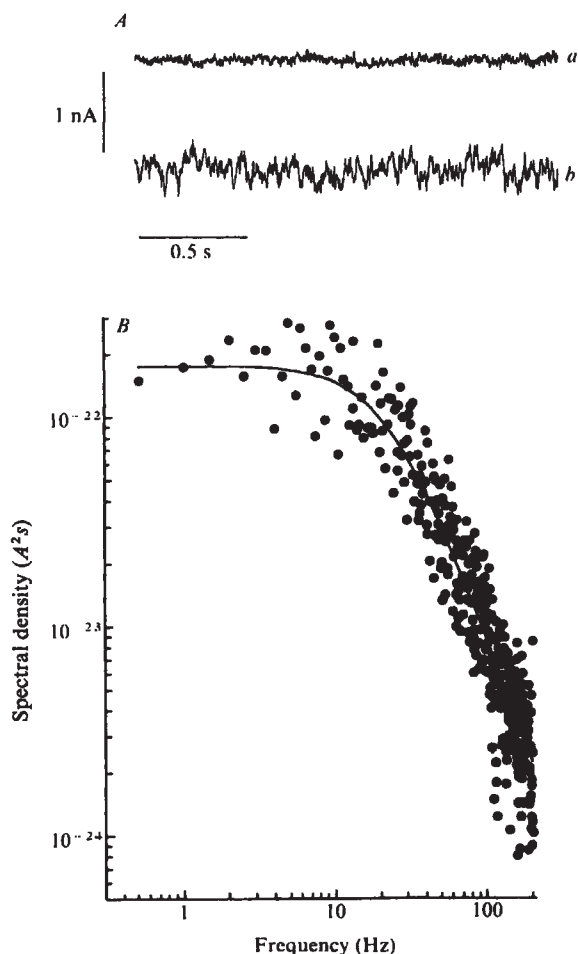


Fig. 1 *A*, Digitised records of current noise before (*a*) and during (*b*) an ACh-induced current of 2.7 nA amplitude. Records are bandpass filtered (0.3–250 Hz, 24 dB per octave rolloff). Temperature 23 °C, holding potential, –100 mV. *B*, Power spectrum of ACh-induced current noise. Temperature 23 °C, holding potential –80 mV, mean ACh-evoked current (μ) 3.5 nA. Clamp currents were recorded on magnetic tape by an f.m. tape recorder (Ampex PR2200, 5 kHz cutoff). The input to one channel was bandpass filtered (0.3 Hz–3 kHz, 24 dB per octave rolloff) and amplified at gains of 350–1,150 mV nA⁻¹. A d.c. record was stored on a second tape channel using gains of 10–100 mV nA⁻¹. Analog data were digitised after low pass filtering (250 Hz cutoff, 24 dB per octave rolloff) to give records of 1.96 s length each containing 2,048 points. Typically, samples of control noise were 5–6 digitised records in length and samples of ACh-evoked noise were 8–10 segments long. The fast Fourier transform was performed on each record and the resulting periodograms averaged for all records in a given noise sample. The spectrum obtained from control noise was subtracted from the ACh noise spectrum. A function of the form given in the text was fitted to the resulting difference spectrum by a least squares method (solid line in Fig. 1*B*). All computations were made on a CDC 3600 computer. For the illustrated spectrum the estimated values of τ and γ were 6.92 ms and 47.9 pS, respectively.

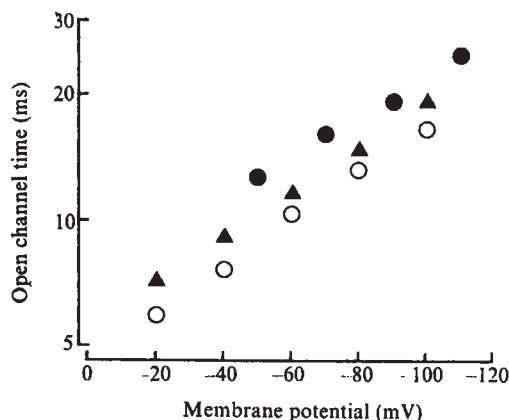


Fig. 2 Semilogarithmic plots of ACh-activated channel lifetime at various membrane potentials in 3 myotubes at 23 °C.

serum reduced myotube ACh sensitivity by about 95%. No obvious change in the conductance of single functional channels was noted after serum treatment. In three cells studied successfully at a holding potential of –70 mV, γ was 44.0 ± 1.7 pS. The channel closing rate was, however, slightly faster; the mean channel lifetime at –70 mV was only 7.2 ± 0.7 ms compared with the value of 12.2 ms found in control myotubes (Mann-Whitney test $P = 0.05$). The mean charge passed per single channel opening in myasthenic serum-treated myotubes (2.2×10^{-14} coulomb at –70 mV and 23 °C) is only 65% of that passed in control myotubes in the same conditions (3.4×10^{-14} coulomb). In view of the small number of myasthenic serum-treated cells studied, this estimate of charge reduction must be viewed with caution. A reduction in the charge passed per open channel is nevertheless consistent with our finding that antibodies raised experimentally against ACh receptors can modify slightly the properties of ACh-activated channels in cultured rat myotubes¹⁰. It is clear, however, that the greatly reduced ACh sensitivity of human muscle cells exposed in this myasthenic serum cannot be explained solely by a modification of functional ACh channel properties. An antibody-induced decrease in the density of membrane ACh receptors resulting from an elevated rate of receptor degradation is probably of more importance in lowering ACh sensitivity¹⁴.

We thank Dr E. H. Lambert and Dr J. Lindstrom for providing the foetal and adult muscle, and Dr S. Heinemann for the cultures. S.B. was supported by a Muscular Dystrophy Associations of America fellowship and R. W. K. by a Canadian MRC fellowship. J. R. was supported in part by a US NSF grant (MPS75-08003).

S. BEVAN*
R. W. KULLBERG†

Neurobiology Department,
The Salk Institute,
P.O. Box 1809,
San Diego, California 92112

Department of Mathematics,
University of California at San Diego,
La Jolla, California 92307

J. RICE

Received 16 January; accepted 21 March 1978.

*Present address: Department of Zoology, University College, London, Gower Street, London WC1, UK.

†Present address: Division of Mathematics and Science, University of Alaska, Anchorage, Alaska 99504.

1. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655–691 (1973).
2. Katz, B. & Miledi, R. *J. Physiol., Lond.* **224**, 665–699 (1972).
3. Colquhoun, D., Dionne, V. E., Stenbach, J. H. & Stevens, C. F. *Nature* **253**, 204–206 (1975).
4. Sachs, F. & Lecar, H. *Biophys. J.* **17**, 129–143 (1977).
5. Lass, Y. & Fischbach, G. D. *Nature* **263**, 150–151 (1976).

6. Dreyer, F., Muller, K. D., Peper, K. & Sterz, R. *Pflugers Arch. ges. Physiol.* **367**, 115-122 (1976).
7. Sakmann, B. *Pflugers Arch. ges. Physiol.* **359**, R89 (1975).
8. Colquhoun, D., Large, W. A. & Rang, H. P. *J. Physiol., Lond.* **266**, 361-395 (1977).
9. Vogt, M. & Dulbecco, R. *Proc. natn. Acad. Sci. U.S.A.* **49**, 171-180 (1963).
10. Heinemann, S., Bevan, S., Kullberg, R., Lindstrom, J. & Rice, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3090-3094 (1977).
11. Brown, K. T. & Fleming, D. G. *Science* **185**, 693-695 (1974).
12. Cull-Candy, S. G., Miledi, R. & Trautmann, A. *Nature* **271**, 74-75 (1978).
13. Neher, E. & Sakmann, B. *J. Physiol., Lond.* **258**, 705-729 (1976).
14. Dreyer, F., Walther, G. & Peper, K. *Pflugers Arch. ges. Physiol.* **366**, 1-9 (1976).
15. Bevan, S., Kullberg, R. W. & Heinemann, S. F. *Nature* **267**, 263-265 (1977).

Colchicine potentiates β -adrenoreceptor-stimulated cyclic AMP in lymphoma cells by an action distal to the receptor

THE cellular cytoskeleton, composed of microtubules and microfilaments, seems to regulate cell surface receptors for a number of membrane-active agents, including lectins¹, immunoglobulins² and polypeptide hormones³. Patching and capping of membrane receptors is controlled by the activity of microfilaments and microtubules, and, in addition, the cytoskeletal elements may be involved in stabilising the fluidity of the cell surface^{4,5}. Colchicine, an inhibitor of microtubular assembly, and presumably, therefore, of microtubule-associated functions, has been used experimentally to study cellular events thought to be regulated by microtubules⁶, and clinically in the treatment of gout and other inflammatory disorders⁶. Evidence from several laboratories indicates that increases in cellular adenosine 3',5'-monophosphate (cyclic AMP) alter the assembly of microtubules⁷⁻¹². It has been suggested that colchicine augments hormone-stimulated cyclic AMP synthesis in target cells for some^{13,14}, but not all^{15,16}, hormones. To study this latter action of colchicine on a well-defined hormonal system, we

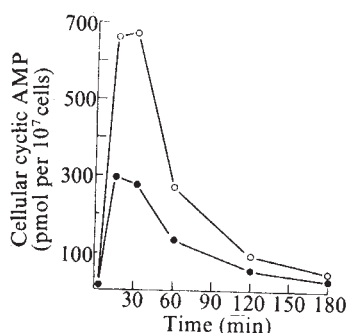


Fig. 1 Time course of cellular cyclic AMP levels in response to isoprenaline in wild-type S49 cells preincubated in the presence (O) and absence (●) of colchicine. Wild-type S49 cells, growing in log phase in suspension culture as described elsewhere^{17,18}, were centrifuged and then resuspended in their growth media, Dulbecco's modified Eagle's medium containing 10% heat-inactivated horse serum. An aliquot of cells was removed and colchicine (Sigma) at a final concentration of 10 μ M was added. Colchicine-treated and control cells were then placed for 1 h in a 37 °C incubator whose atmosphere was 95% air, 5% CO₂. Isoprenaline (1 μ M, final concentration) in ascorbic acid (0.1 mM, final concentration) was added to both batches of cells and aliquots were removed at the indicated times for the determination of cellular cyclic AMP content. The aliquots were spun at 400g for 3 min in a refrigerated centrifuge, the media was aspirated, and the cells were resuspended and then boiled for 3 min in 0.5 ml of 50 mM sodium acetate, pH 4.0, containing 0.2 mM 3-isobutyl-1-methylxanthine (Sigma). Cyclic AMP was determined using a competitive binding protein method¹⁹. The data are expressed as cyclic AMP per 10⁷ viable cells. Cell counts were done with a haemocytometer and viability (>90%) was assessed by the exclusion of the supravital stain erythrosine B.

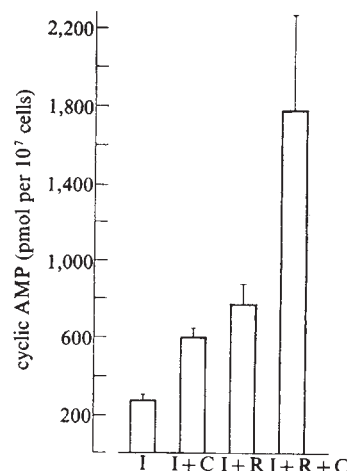


Fig. 2 Effect of isoprenaline (I) alone or with Ro 20-1724 (R), colchicine (C) or both (I+R+C) on cyclic AMP levels in S49 cells. Logarithmically growing cells were counted, spun and then resuspended in culture media as described in Fig. 1. The cells were then incubated for 30 min at 37 °C in the presence of 1% dimethylsulphoxide alone or containing 50 μ M colchicine and/or 100 μ M Ro 20-1724 (a potent phosphodiesterase inhibitor in S49 cells¹⁹). Cells were then spun and resuspended in Dulbecco's modified Eagle's medium altered so that 20 mM HEPES was substituted for NaHCO₃ and 0.1% bovine serum albumin was substituted for horse serum (DMEB medium). Isoprenaline (1 μ M, final concentration) in ascorbic acid (1 mM, final concentration) was added, and samples were incubated for 6 min at 37 °C. The reactions were stopped by adding 0.5-ml aliquots of cells to 0.05 ml 300 mM acetic acid and then heating immediately at 85 °C for 3 min. Cyclic AMP was measured as described in Fig. 1 and compared to cyclic AMP standards suspended in media that had been acidified and boiled in a manner identical to the experimental samples. The data are the mean $\pm$ s.d. of three experiments.

have examined the augmentation by colchicine of β -adrenergic amine-stimulated cyclic AMP levels in S49 mouse lymphoma cells, a model system for studying the synthesis of cyclic AMP and the actions of β -adrenergic amines^{17,18}. We report here that colchicine potentiates β -adrenergic-stimulated cyclic AMP accumulation in S49 cells and that this potentiation results from events occurring beyond the interaction of β -adrenergic agonists with β -adrenoreceptors.

S49 cells respond to the β -adrenergic agonist isoprenaline with a marked increase in cellular cyclic AMP (Fig. 1). This increase occurs rapidly but only transiently, presumably because of the refractoriness that these¹⁹ and many other cells^{18,20} demonstrate in response to adrenergic agonists. When S49 cells are treated with colchicine before the addition of isoprenaline, they show a potentiation in their response to the adrenergic agonist, although the temporal pattern of the response parallels that of cells incubated with isoprenaline alone. Thus, colchicine does not seem to block refractoriness in these cells. The enhanced cyclic AMP levels following colchicine treatment do not result from decreased cellular extrusion of the nucleotide, because determination of cellular cyclic AMP or cyclic AMP in media plus cells that have been incubated with colchicine yield results similar to those shown in Fig. 1.

Results from several experiments indicate that enhancement of isoprenaline stimulated cyclic AMP by colchicine is two to sixfold, is potentiated by phosphodiesterase inhibitors (Fig. 2), and does not occur when the optically altered derivative lumicolchicine is used instead of colchicine. In addition to the results of experiments in which a phosphodiesterase inhibitor was used (Fig. 2), direct measurements of cyclic nucleotide phosphodiesterase activity indicate that colchicine's effect cannot be attributed to a decrease in cellular phosphodiesterase activity (data not

shown). From this, we conclude that colchicine increases cyclic AMP accumulation in response to isoprenaline by increasing the activity of the β -adrenoreceptor-adenylate cyclase system.

To test whether colchicine's effect results from a change in binding properties of the β -adrenoreceptors, we studied these receptors using intact S49 cells in the presence and absence of colchicine. For these studies, we used the radio-labelled β -adrenergic antagonist ^3H -dihydroalprenolol (DHA). This radioligand, which has been used in identifying and quantifying β -adrenoreceptors in a variety of intact²¹⁻²³ and disrupted²⁴⁻²⁶ cell preparations, has enabled us to study cellular cyclic AMP accumulation and binding to β -adrenoreceptors in similar experimental conditions. Binding studies using DHA and intact S49 cells carried out as described in the legend to Fig. 3 fulfill the criteria expected for interaction at the β -adrenoreceptor^{18,27}: specific binding of DHA achieves equilibrium rapidly (<3 min at 37°C), is fully reversible ($t_{1/2} = 5$ min), and is competed for stereoselectively by β -adrenergic agonists and antagonists. The rank order of potency for such agents in the binding studies parallels that observed for stimulating or inhibition of cellular cyclic AMP accumulation by the same agents in these cells. Equilibrium binding studies demonstrate a single class of binding sites, and the dissociation constant (K_d) for DHA estimated from both kinetic and equilibrium studies are in good agreement.

After the cells had been incubated with colchicine (5×10^{-5} M) for 30–60 min, a time when maximal potenti-

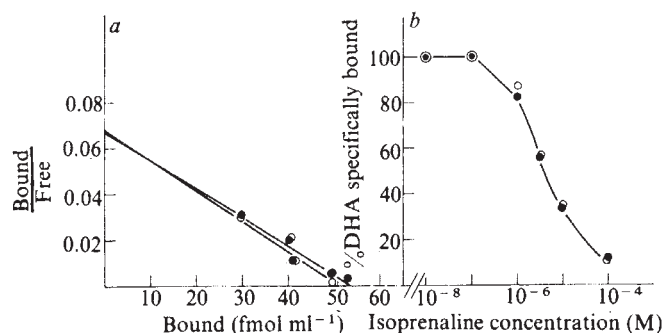


Fig. 3 Effect of colchicine on DHA binding of S49 cells. Logarithmically growing cells were incubated for 30 min at 37°C in the presence (O) or absence (●) of $50 \mu\text{M}$ colchicine. The cells were then centrifuged, washed once in DMEB buffer, recentrifuged, and resuspended in DMEB buffer. Binding assays were run by incubating cells at 37°C for 5 min in a volume of 0.25 ml DMEB buffer containing 13,000–210,000 c.p.m. ^3H -dihydroalprenolol (New England Nuclear, 54 c.p.m. fmol^{-1}), 1 mM ascorbic acid or 1 mM ascorbic acid and variable concentrations of (–)-isoprenaline (b) or ascorbic acid and (–)-alprenolol (3×10^{-7} M, final concentration)(a). At each concentration of DHA, cells were incubated in the absence and presence of (–)-alprenolol. Binding reactions were ended by adding 3 ml of 5 mM ice-cold buffer containing potassium phosphate, pH 7.0, 1 mM MgSO_4 , and 0.01 mM ($\pm$)-propranolol and then filtering immediately over glass fibre filters (Gelman A/E). The filters were rinsed with 15 ml of cold buffer (5 mM potassium phosphate, pH 7.0, 1 mM MgSO_4), dried by vacuum and then counted in a liquid scintillation system (efficiency 50%). Propranolol was included in the initial stop buffer to eliminate nonspecific retention of counts on the filters, and, thus, to equalise samples incubated with ligand alone or radioligand plus competitors of binding. All samples were run in duplicate and these differed from each other by less than 10%. Nonspecific binding, the number of counts bound in the presence of 3×10^{-7} M (–)-alprenolol, was less than 30% of the total counts bound except at the highest concentrations of DHA (15 nM), at which nonspecific counts were 50% of total counts bound. a, A Scatchard plot of specific binding of DHA to cells incubated in the absence (●) and presence (O) of colchicine. b, The inhibition of specifically bound counts by increasing concentrations of (–)-isoprenaline in an experiment in which the DHA concentration was 4 nM.

ation of isoprenaline-stimulated cyclic AMP was observed, we determined the number of DHA binding sites and the affinity of these sites for DHA and isoprenaline. The K_d (control, 0.9 ± 0.3 nM; colchicine incubated, 1.2 ± 0.6 nM, mean $\pm$ s.d.) and number (control, $1,280 \pm 73$; colchicine-incubated, $1,386 \pm 66$) of DHA binding sites are not significantly altered by colchicine. A representative Scatchard plot of data from one experiment is shown in Fig. 3. Moreover, studies in which isoprenaline competes with DHA for binding sites on the cells indicate that colchicine does not change the molar potency of isoprenaline in competing with DHA (Fig. 3). This lack of effect of colchicine in the β -adrenoreceptor binding studies indicates that colchicine-induced stimulation of cellular cyclic AMP results from an effect that occurs beyond the interaction of the β -adrenergic agonist with its receptor.

The absence of an effect of colchicine on binding of a hormone to its receptor on the plasma membrane has also been reported for insulin and growth-hormone binding³, but in S49 cells this lack of effect on the β -adrenoreceptor occurs in a setting of increased β -adrenergic stimulation of cyclic AMP. Our results therefore apparently localise the action of colchicine to events that regulate expression of β -adrenergic-stimulated adenylylase activity. The mechanism whereby hormone-receptor complexes activate adenylylase is being investigated in several laboratories^{18,28-31}, but none of the proposed models for cyclase activation has been definitively confirmed³². For this reason, we cannot define precisely the mechanism(s) by which colchicine potentiates isoprenaline activation of cyclase. It has been suggested³ that colchicine may interact with plasma membrane tubulin or other membrane components. However, our findings in S49 cells (data not shown) as well as the results mentioned but not shown by Rudolph *et al.*¹⁴ in polymorphonuclear leukocytes indicate that colchicine does not activate adenylylase activity in plasma membrane preparations *in vitro*. In both systems, intact cells are required for colchicine-induced potentiation of hormonal agonists, and hence, its interaction with cytosolic tubulin may be responsible for this potentiation.

This work was supported by NIH grants GM 00001, GM 16496, HL 06285, NSF grant PCM 14397 and money from the University of California. We thank Michele Sanda for technical assistance. P.A.I. is an Established Investigator, American Heart Association. RO 20-1724 was supplied by Dr H. Shephard, Roche.

PAUL A. INSEL
MICHAEL S. KENNEDY

Division of Clinical Pharmacology,
Departments of Medicine and Pharmacology,
Cardiovascular Research Institute,
University of California,
San Francisco, California 94143

Received 3 February; accepted 3 April 1978.

- Ukena, T. E., Borysenko, J. Z., Karnovsky, M. J. & Berlin, R. D. *J. Cell Biol.* **61**, 70–82 (1974).
- Taylor, R. B., Duffus, W. P. H., Raff, M. C. & de Petris, S. *Nature new Biol.* **233**, 225–229 (1971).
- Van Obberghen, E., De Meyts, P. & Roth, J. *J. biol. Chem.* **251**, 6844–6851 (1976).
- Edelman, G. M. *Science* **192**, 218–226 (1976).
- Stephens, R. E. & Edds, K. T. *Physiol. Rev.* **56**, 218–226 (1976).
- Malawista, S. E. *Arthritis Rheum.* **56**, 709–711 (1976).
- Gillespie, E. *Ann. N.Y. Acad. Sci.* **253**, 771–779 (1975).
- Dipasquale, A. M., McGuire, J., Moellmann, G. & Wasserman, S. J. *J. Cell Biol.* **71**, 735–748 (1976).
- Weissmann, G., Goldstein, I., Hoffstein, S. & Tsung, P. *Ann. N.Y. Acad. Sci.* **253**, 750–762 (1975).
- Willingham, M. C. & Pastan, I. *J. Cell Biol.* **67**, 148–159 (1975).
- Rubin, R. W. & Weiss, G. D. *J. Cell Biol.* **64**, 42–53 (1975).
- Fuck, T. T. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4491–4495 (1977).
- Gemsa, D., Steggemann, L., Till, G. & Resch, K. *J. Immun.* **119**, 524–529 (1977).
- Rudolph, S. A., Greengard, P. & Malawista, S. E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3404–3408 (1977).
- Iyengar, R., Lepper, K. G. & Mailman, D. S. *J. supramolec. Struct.* **5**, 521–530 (1976).
- Kaliner, M. *J. clin. Invest.* **60**, 951–959 (1977).
- Coffino, P. *et al. Recent Prog. Horm. Res.* **32**, 669–682 (1976).

18. Maguire, M. E., Ross, E. M. & Gilman, A. G. in *Advances in Cyclic Nucleotide Research* 8 (eds Greengard P. & Robieson, G. A.) 1-84 (Raven, New York, 1977).
19. Shear, M., Insel, P. A., Melmon, K. L. & Coffino, P. J. *biol. Chem.* **251**, 7572-7576 (1976).
20. Insel, P. A. in *International Rev. of Biochemistry, series II, Biochemistry and Mode of Action of Hormones* (ed. Rickenberg, H. V.) (University Park Press, Baltimore, in the press).
21. Mukherjee, C. & Lefkowitz, R. J. *Life Sci.* **19**, 1897-1906 (1976).
22. Tallman, J. F., Smith, C. C. & Henneberry, R. C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1564-1568 (1975).
23. Sheppard, J. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1091-1094 (1977).
24. Mukherjee, C., Caron, M. G., Coverstone, M. & Lefkowitz, R. J. *J. biol. Chem.* **251**, 4869-4876 (1975).
25. Alexander, R. W., Williams, L. T. & Lefkowitz, R. J. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1564-1568 (1975).
26. Zatz, M., Kebabian, J. W., Romero, J. A., Lefkowitz, R. J. & Axelrod, J. *J. Pharmac. exp. Ther.* **196**, 714-722 (1976).
27. Wolfe, B. B., Harden, T. K. & Molinoff, P. B. *Ann. Rev. Pharmac.* **17**, 575-604 (1977).
28. Helmreich, E. J. M., Zenner, H. P., Pfeuffer, T. & Cori, C. F. in *Current Topics in Cellular Regulation* (eds Horecker, B. L. & Stadtman, E. R.) **10**, 41-87 (Academic, New York, 1976).
29. Levitzki, A. *Biochem. biophys. Res. Commun.* **74**, 1154-1159 (1977).
30. Rendell, M. S., Rodbell, M. & Berman, M. *J. biol. Chem.* **252**, 7909-7912 (1977).
31. Cuatrecasas, P. & Hollenberg, M. D. *Adv. Protein Chem.* **30**, 251-451 (1976).
32. Greaves, M. F. *Nature* **265**, 681-683 (1977).

Release of histamine from rat mast cells by acetylcholine

PARASYMPATHETIC stimulation using either acetylcholine or carbamylcholine at nanomolar to picomolar concentrations enhances antigen-induced release of histamine from various IgE sensitised tissues (human lungs¹, human nasal polyps²). The enhancement induced by cholinomimetic compounds is blocked by pretreatment with atropine, so is probably affected through a muscarinic receptor³. However, enhancement of histamine release by cholinergic treatment has not been observed from peripheral human leukocytes sensitised with human IgE (ref. 3). This discrepancy, as well as the lack of information on the effect of acetylcholine on histamine-storing cells, have prompted us to investigate the action of acetylcholine on non-evoked histamine release from isolated purified rat mast cells. To our knowledge, only Kiernan⁴ has studied the effects of acetylcholine on isolated mast cells, and failed to substantiate any significant morphological change. We present here the results of a study of the effects of acetylcholine on the morphology and on histamine release of rat mast cells, and show that low concentrations of acetylcholine are sufficient to cause histamine release.

Mast cells from the abdominal and thoracic cavities of male Wistar albino rats (weight 200-400 g) were collected according to the method of Thon and Uvnäs⁵ and, after gradient centrifugation in Ficoll (Pharmacia), suspended in a medium containing NaCl 145 mM, CaCl₂ 0.9 mM, KCl 2.4 mM, glucose 0.1% and human serum albumin 0.1% adjusted to pH 7.4 with 10% (v/v) Sørensen phosphate buffer. Duplicate aliquots (50 µl) of cell suspension were added to test tubes containing various concentrations of acetylcholine (acetylcholine chloride, BDH) dissolved in medium as above to a final volume of 2 ml. Previous experiments had shown that acetylcholine was quantitatively recovered after incubation for 1 h in a medium containing mast cells. The samples were incubated at 37 °C in a metabolic shaker for 10 min. The reaction was stopped by chilling the tubes in an ice-water bath. Cells were then separated from the medium by centrifugation (400g for 5 min) at 4 °C. Histamine was measured fluorimetrically using the method of Shore *et al.*⁶ as modified by Kremzner and Wilson⁷. In the supernatants, o-phthalaldehyde was added directly to the samples after alkalisation. The same procedure was used for the cells, after extraction with HCl 0.1 M using the method of Bergendorff and Uvnäs⁸. In some experiments the histamine assay was carried out in the same sample both through the fluorimetric assay and through the bioassay in the presence of atropine 3 × 10⁻⁷ g ml⁻¹ and methysergide 10⁻⁸ ml⁻¹

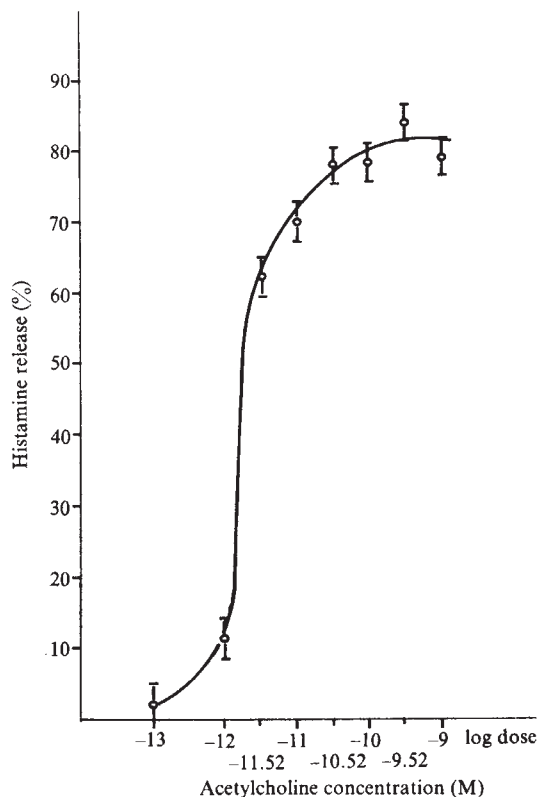
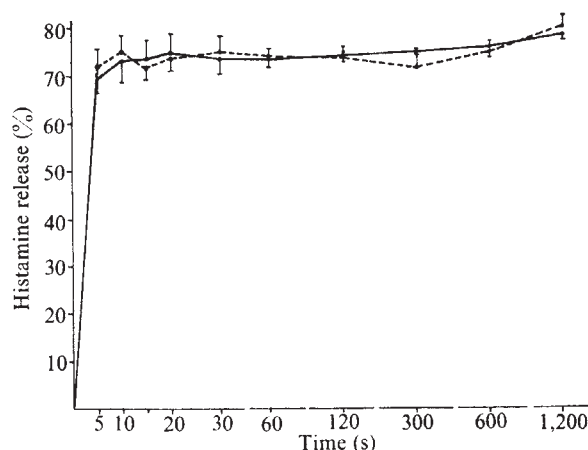


Fig. 1 Log dose-response effect of acetylcholine on histamine release from mast cells. Each point represents the mean $\pm$ s.e.m. of six experiments.

(ref. 9). The authenticity of histamine in control and treated samples was demonstrated either by recording the excitation and fluorescence spectra or through the bioassay according to a 2 by 2 design¹⁰. At any given concentration acetylcholine did not interfere with histamine assay. Histamine release (supernatant histamine) was expressed as a percentage of the total present in the cells plus supernatants. Spontaneous histamine release ranged between 1 and 8% and was subtracted from all values.

Rat mast cells purified to > 80% released histamine when exposed to very low concentrations of acetylcholine (10⁻¹³ to 10⁻⁹ M) (Fig. 1). Time course studies demonstrated that the release of histamine evoked by acetylcholine was similar to that induced by compound 48/80 (Fig. 2). The process was blocked by low temperature, by glucose depriva-

Fig. 2 Time course of the histamine release from mast cells induced by 10⁻¹⁰ M acetylcholine (—) and 10⁻⁶ g ml⁻¹ 48/80 (---). Each point represents the mean $\pm$ s.e.m. of six experiments.



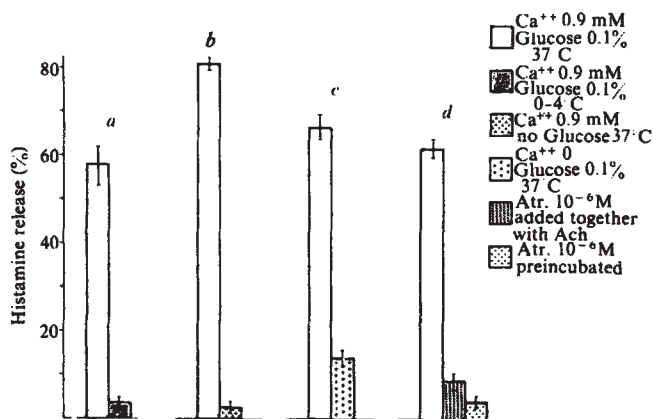
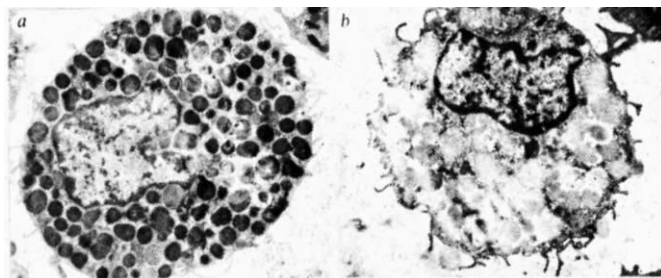


Fig. 3 Effect on histamine release of *a*, temperature. Mast cells induced by 10^{-10} M acetylcholine (—) and 10^{-6} g choline, then at 37°C for 10 min with the releasing agent¹⁴. *b*, Glucose deprivation. Cells were incubated in a glucose-free medium. *c*, Withdrawal of extracellular calcium. Cells were incubated in a calcium-free medium. *d*, Atropine. In the first group of experiments, cells were preincubated with atropine for 15 min, then centrifuged and resuspended in the solution containing acetylcholine; in the second group of experiments, both drugs were added simultaneously to the cell suspension. Atropine up to 10^{-4} M was devoid of any histamine releasing activity. In each case, 10^{-10} M acetylcholine was used, six experiments were carried out, and the mean $\pm$ s.e.m. ($P>0.01$) is indicated.

tion, by withdrawal of extracellular calcium, and by atropine. Atropine caused blocking either when added together with acetylcholine, or when the cells were preincubated and washed before exposure to acetylcholine (Fig. 3).

Electron microscopy of a control cell is shown in Fig. 4a. The granules are present as homogeneous, dense-staining cytoplasmic inclusions, surrounded by their granule membranes. The cell membrane is associated with a few microvilli, but it is otherwise relatively smooth in contour. After treatment with acetylcholine, electron microscopy reveals granule changes (Fig. 4b), including the enlargement of the space between the granule and its membrane, swelling, and a reduction in its electron density and homogeneity. Some granules are discharged from the cell, but most remain within the cell, lying in a network of cavities that have been shown to communicate with the extracellular space. Similar features have been reported in rat peritoneal

Fig. 4 Isolated mast cells were incubated for 10 min at 37°C in the medium previously described with or without acetylcholine (10^{-10} M), then collected by centrifugation and fixed by the addition of a double volume of cold glutaraldehyde (3% in 0.1 M Na-cacodylate buffer, pH 7.3; $0-4^{\circ}\text{C}$ for 15 min; room temperature for 60 min). Post-fixation was in 1% osmium tetroxide (OsO_4) in 0.1 M Na-cacodylate buffer at room temperature, according to the procedure described by Lawson *et al.*¹¹. Electron micrographs: *a*, control mast cells ($\times 10,500$); *b*, treated mast cells ($\times 10,500$).



mast cells exposed to 48/80 (type 3 and type 2 cells, according to Krüger¹²).

Our results, showing that acetylcholine can affect rat mast cells by evoking a sequential exocytosis, suggest that it would be worthwhile to study the relationship between cholinergic activity and the secretory response of mast cells.

ROBERTO FANTOZZI

EMANUELA MASINI

PATRIZIO BLANDINA

PIER FRANCESCO MANNAIONI

TATIANA BANI-SACCHI

Department of Pharmacology,

Department of Histology,

University of Florence,

Viale G. B. Morgagni, 50134-Firenze, Italy

Received 14 August 1977; accepted 20 February 1978.

1. Kaliner, M., Orange, R. P. & Austen, K. F. *J. exp. Med.* **136**, 556-567 (1972).
2. Kaliner, M., Wassermann, S. I. & Austen, K. F. *New Engl. J. Med.* **239**, 277-281 (1973).
3. Kaliner, M. & Austen, K. F. *A. Rev. Pharmac.* **15**, 177-189 (1975).
4. Kiernan, J. A. *Experientia* **28**, 653-655 (1972).
5. Thon, I. L. & Uvnäs, B. *Acta physiol. scand.* **71**, 303-315 (1967).
6. Shore, P. A., Burkhalter, A. & Cohn, V. H. *J. Pharmac. exp. Ther.* **127**, 182-186 (1959).
7. Kremzner, L. T. & Wilson, I. B. *Biochim. biophys. Acta* **50**, 364-367 (1961).
8. Bergendorff, A. & Uvnäs, B. *Acta physiol. scand.* **84**, 320-331 (1972).
9. Giotti, A., Guidotti, A., Mannaioni, P. F. & Zilletti, L. *J. Physiol., Lond.* **184**, 924-941 (1966).
10. Schild, H. O. *J. Physiol., Lond.* **101**, 115-130 (1942).
11. Lawson, D., Raff, M. C., Gomperts, B. D., Fewtrell, C. & Gitula, N. B. *J. Cell Biol.* **72**, 242-259 (1977).
12. Krüger, P. G. *Int. Archs Allerg. appl. Immun.* **51**, 608-626 (1976).
13. Code, C. F. *New Engl. J. Med.* **296**, 1459-1462 (1977).
14. Johnson, A. R. & Moran, N. C. *J. Pharmac. exp. Ther.* **189**, 221-234 (1974).

Corrigenda

In the letter 'Unintegrated ribosomal genes and their relation to position effect variegation in *Drosophila melanogaster*' by C. I. Zuchowski-Berg, *Nature* **271**, 60, the author's address should read: Division of Cell and Molecular Biology, SUNY at Buffalo, New York 14214. The authors present address is: Department of Biology, Princeton University, Princeton, New Jersey 08540.

In the letter 'Selection effects in redshift distribution of QSOs' by D. Basu, *Nature* **273**, 310, the journal in ref. 2 should be *Astrophys. Lett.*

Nature Index and Binders

The complete Index for 1977 is available, price £2.50 (U.K.), US\$5.00 (Rest of World). Copies of the 1976 index are still on sale, prices as above.

Binders for the journal are also available at £8.00 (UK), US\$16.00 (Rest of World) for three: a year of *Nature* fits into three binders.

All prices include postage. Payment may be made in any currency at the prevailing exchange rate. Orders should be sent, accompanied by remittance, to Macmillan Journals Ltd, Brunel Road, Basingstoke, Hampshire, RG21 2XS, England.

reviews

Alternatives to growth-economy

Eric Ashby

Alternatives to Growth: A Search for Sustainable Futures. Vol. 1. Edited by Dennis L. Meadows. Pp. 405. (Ballinger: Cambridge, Massachusetts; Wiley: Chichester, UK, 1978.) £10.40; \$18.50 in UK, Europe and Africa.

THIS book is edited by Dennis Meadows, one of the authors of the notorious *Limits to Growth*. But he is at pains to point out in the introduction that *Alternatives to Growth*, Volume 1, and the four volumes which are to follow, are not intended to justify the notion of limits to growth. True, several of the eighteen contributors to this volume are committed to belief in a steady-state economy for the industrial nations. But in this book they come out of their pulpits (or off their soap boxes) to sit at the discussion table, to consider how, if there are to be alternatives to the present growth-economy, these alternatives might work. The book deserves a welcome as a serious attempt to put bridges across the gap between those who pontificate about our present perplexities and those who have to make decisions about how to deal with the perplexities.

Like all collections of papers presented at a conference (the first of five conferences to be held on the initiative of a Texas businessman, Mr George Mitchell), the contributions are uneven in quality. They are arranged under four sections: nutrition and energy in the steady state; economic alternatives in an age of limits; the politics of equity and social progress in a finite world; life-styles and social norms for a sustainable state. The book must be judged by its declared intention: 'to stimulate development of new information useful to decisionmakers in the industrialized West, as they work to decide what must be done *by them*, now, and *within their own countries*'. (Italics are the editor's.)

Some of the essays fall short of this intention. Thus, an account of the enterprises of the New Alchemy Institute, which encourages the design of "arks"—complete with solar energy, fish-farm attached, wind-powered, and growing their own food—is fun to read, but as "information useful to decisionmakers" it is not likely to relieve the anxieties of the Secretary of State for Energy. And, to take another example, from an essay on settlements and social stability, although there is wisdom to be learnt from a study

of the structure of primitive societies, the difficulties of applying this wisdom to the city government of Detroit or Birmingham deter the most enterprising decision-maker from using this sort of information. So, for the purpose for which these essays were written, some of them are disappointing.

But there are other essays which do confirm what is now becoming a widespread conviction among thoughtful people; namely, that issues like energy policy, the conservation of oil and other non-renewable resources, the investment of capital (or failure to invest capital) in labour-saving machinery—these are not primarily technological problems; they are moral and political problems, and such problems do not have technological solutions. Thus, one essay (in my view the best in the book) describes the symptoms of the present sickness in our economy; a constantly increasing proportion of the GNP being spent on preventing breakdowns in society, on "mediating conflicts, controlling crime, footing the bill for all social costs generated by the "externalities" of production and consumption . . . and generally trying to maintain 'social homeostasis' . . .".

The most serious symptom is that some of the significant externalities are still left out of account. Not only are there the familiar externalities of pollution (which are being taken into account); there are also social externalities—family relationships, community cohesion, and the network of social sanctions. These, too, are being damaged as a consequence of the nation's dedication to economic growth.

Of course, the growth of the industrial West has been an unprecedented success story; no-one but a fool hankers after the barbarities of early Victorian England—the subjection of women, the exploitation of children, the poor nutrition, the prevalence of disease, the indifference to the environment. But the time has come to count the cost of these successes. Medical science has dramatically reduced mortality; yes, but it has also generated pockets of over-population, and a problem of care for the aged. High technology has changed the dimensions of space, time and memory; but it has also made whole cities vulnerable to sabotage or 'acts of God' (and, as one of the essayists remarks, to make it a condition that no acts of God can be permitted is to commit the sin of

hubris). And the computer which books air flights and issues the bank statement and receives the gas bill can also be applied to the invasion of human privacy and the erosion of individual freedom.

Communication has been revolutionised; but the revolution has burdened people with such an overload of information that their judgement is impaired. A widespread provision of basic needs—a car in nearly every garage, a television set in nearly every living room, a refrigerator in nearly every kitchen: these have not been a source of satisfaction; they have aroused expectations, sometimes to be bitterly disappointed, among the minority (in some countries) and the majority (in others) who do not possess these amenities. The greatly enlarged powers of choice open to affluent man confront him with a fresh dilemma—some of the choices will have nasty second order effects he hasn't anticipated; and the whole problem of managing choice when it is so diverse has not been solved.

Wealth has expanded beyond the dreams of our grandparents; it has also widened the gap between the 'haves' and the 'have-nots', and the wider the gap the greater the resentment. There is some sense, therefore, in the suggestion described in the essay on the "coming economic transition", made by a body calling itself Environmentalists for Full Employment: that there should be an employment-impact-statement required for a capital development project, in a nation which has a chronic unemployment rate of 7%.

The essay on limits to energy conversion contains one very perceptive remark. It is about the curtailment of personal freedom which people may be prepared to accept in order to rid themselves of fear about sabotage of nuclear plants (and, for that matter, other highly centralized services), and in order to diminish anxieties about ups and downs in the management of the economy. This acquiescence (the writer was referring particularly to nuclear power) "could bring about precisely those political changes that our costly military nuclear deterrent was intended to prevent".

The most persuasive theme in the book is the reiterated fear that the social and ethical consequences of high technology may be beyond the capacity of a free

market economy, and even of a pluralistic democratic government, to control. The challenge of the book is to ask: If this fear turns out to be justified, how do you order your priorities between security, equity, freedom, and efficiency?

The remedies suggested, as a guide to the decision-maker, are not (most of them) original. Decentralisation is advocated, of course—into small units where local leaders can work for local consensus; so is a negative income tax, to pacify a population which might otherwise revolt during the painful journey from a growth-economy to a steady-state economy; and so is rationing, as a control for scarce resources, rather than reliance on the market price. One attractive idea is to issue a petrol ration to every person over the age of 18, whether he or she has a car or motor bike

or not. The total ration would represent the nation's decision as to how much petrol may be used for private consumption. Owners of ration tickets could then sell them, or the state could buy them back at an agreed price. If the ration is adequate, the temptation to have a black market would be minimised.

This volume, and the four more which are to come, may not offer blueprints to politicians. But the first volume is a healthy reminder to decision-makers that although technological fixes may keep the democratic West going for some time yet, in the long term the problems they—and the people on whose behalf they make decisions—have to solve are not technological; they are moral. □

Lord Ashby is Chancellor of the Queen's University, Belfast, and Fellow of Clare College, Cambridge, UK.

Accounting for cognitive psychology

The Psychology of Cognition. By Gillian Cohen. Pp. 241. (Academic: London, New York and San Francisco, 1977.) £7.50; \$15.65. *The Process of Cognition.* By Arthur L. Blumenthal. Pp. 230. (Prentice-Hall International: Englewood Cliffs, New Jersey and Hemel Hempstead, UK, 1977.) £7.65; \$10.95.

THE goal of cognitive psychology is to explain how the mind works. Its method is experimental, and its essence can best be illustrated by two of the many phenomena that it has uncovered. First, consider an experiment in which subjects are shown an array of letters that is too large to be reported in its entirety after a single brief exposure. If, immediately after such a presentation, the subjects are instructed to describe only part of the array, then they perform very well regardless of the particular part to which their attention is directed. It seems that there is a sensory buffer that briefly prolongs the effective duration of stimuli. Second, experiments have shown that if subjects are prevented from rehearsing a few digits that they have to remember, then their ability to recall them declines exponentially over time. What is to be made of such phenomena? Cognitive psychologists attempt to account for them in small-scale models cast in the form of flow-chart diagrams in which information flows between boxes with such labels as "sensory buffer" and "short term memory". There can be little doubt that the enterprise is scientific, but it has so far failed to yield a comprehensive account of human mentality. These two books, one by an English psychologist, Gillian Cohen, and the other by an American psychologist, Arthur Blumenthal, attempt in their different ways to deal with this problem.

The Psychology of Cognition concen-

trates on language, thinking and problem solving. Each chapter describes some experimental results, outlines several theoretical accounts of them, and then—in the best spirit of English compromise—comes to no very definite conclusion about which explanation is best. After seven such chapters even the most dedicated cognitive psychologist is likely to be feeling rather pessimistic. The one certainty is that the discipline will not make progress if it continues to be practised in the same way. Cohen turns to computers, but here she adopts a different expository strategy. She discusses their use only in simulating human behaviour; she ignores the rising discipline of Artificial Intelligence; and she overlooks the role of programs in developing general psychological theories, a very different matter from simulation in that the program is not even supposed to exhibit all the properties of the process that is modelled. Cohen accordingly ends her book with what might reasonably have been its starting point, the current ills of cognitive psychology. She offers no quick or quack remedies, but just a familiar plea that psychologists should aim for a compromise with physiologists. She argues, not for an outright reduction of psychological phenomena to physiological explanations, but for "the cultivation of greater methodological eclecticism". This prescription is presumably fulfilled by the studies of the functional asymmetries between the two cerebral hemispheres, research to which Cohen has contributed, and to which she devotes a chapter of her book. Unfortunately, it has yet to be shown how the localisation of a mental ability to one hemisphere or the other contributes to an understanding of its underlying mechanism.

The Process of Cognition presents a very different diagnosis of cognitive psychology. Blumenthal has stood back from the details of current controversies and written a work that is imbued with an historical perspective and sense of direc-

tion. He has eschewed the conventional chapter headings of a text, and based his book on the temporal organisation of experience. He starts with the rapid sensory integrations that yield the perceived flow of events, and works his way through to the mechanisms that give rise to the stability and continuity of mental life. His procedure is essentially inductive. Each chapter begins with a pair of related generalisations, reviews some of their empirical consequences, and closes with their recapitulation. Here, for example, is the generalisation that subsumes the existence of a sensory buffer: "Preattentive buffer delays are delays of inputs to experience at a simple or unstructured level of representation prior to incorporation into central attentional integrations . . .". After seven such chapters even the most benighted cognitive psychologist is likely to be feeling mildly optimistic—perhaps the discipline is more unified than one had supposed. In the interests of pedagogy, however, Blumenthal has deliberately ignored inconvenient evidence. He is right to do so: any sort of uniformity is better than being overwhelmed by a chaos of data. Unfortunately, he makes no attempt to move beyond his generalisations to an explanatory theory; and it remains unclear whether he regards such a step as ultimately necessary.

Why is it so difficult to understand the mind? One answer is that psychology is a new science and needs time to develop better methods. The claim is false. Before the Dark Ages of Behaviorism set in, psychologists spent several decades studying mental mechanisms. The two exemplary phenomena described at the outset of this review were, as Blumenthal points out, originally reported in 1879 and 1895, respectively, though most psychologists probably regard them as recent discoveries. Another view is that psychologists have been too stupid to make progress: any competent physicist could 'get up' the subject in a couple of weeks, and clear up the mess in a couple more. There is no answer to this argument other than to say: come and try; you will be welcome, as any of the physicists, chemists, or engineers, already doing psychology will testify. Perhaps the task is really impossible: the mind can no more elucidate itself than a cabbage can. If so, one can only retort that nature has made the exercise seem damnably and tantalisingly feasible. On the supposition that genuine progress can be made, then perhaps the real problem may be one to which Cohen alludes, the stranglehold of the experimental tradition. One experiment is worth a thousand theories in psychology. Until this imbalance is redressed, students of the mind are unlikely to see very deeply into its workings.

P. N. Johnson-Laird

P. N. Johnson-Laird is Reader in Experimental Psychology at the University of Sussex, UK.

Climatic change

Climatic Change. Edited by John Gribbin. Pp. 292. (Cambridge University Press: Cambridge and London, 1978.) Hardback £17.50; paperback £6.50.

TWELVE scientists on the cutting-edge of climatic research have contributed to the five sections of this book: studying past climates, the Earth's heat budget, astronomical influences, modelling changing climates, and climate and man. In general, they assume that their particular field of study is not well understood beyond their own circles, so that each contribution is well suited to the general reader and the specialist in complementary fields of climatic research. As glacial geology provided the first and most convincing evidence for that most important phenomenon of climatic change (the ice ages), it is unfortunate that a section on the current state of glaciology and glacial geology is absent. Even so, the book provides the most complete survey to date of the results of studies by major disciplines on climatic change.

D. H. Tarling reviews the geological-geophysical framework of ice ages. Many of the hypotheses that attempt to explain the Quaternary ice ages cannot be applied to earlier glaciations, so there is a tendency to dismiss these earlier events as nothing but polar ice caps that migrated across former continents drifting through the Earth's rotational axis. Tarling points out, however, that these ice caps waxed and waned with enough regularity and intensity to constitute ice epochs, if not ice ages. He also notes that no single mechanism or synthesis of mechanisms can yet account for these glaciations, as the external boundary conditions were different for each ice epoch. These conditions are the galactic, orbital, palaeomagnetic, palaeogeographical, atmospheric, and tectonic states for each ice epoch. He does not mention, however, the internal boundary conditions that are responsible for glacial surges. These boundary conditions are fundamentally different from and largely independent of the external boundary conditions discussed by Tarling, and surges of ice caps might alter the geographical and atmospheric external boundary conditions enough to trigger ice ages.

Palaeobotany in relation to climatic change is discussed in an informative contribution by T. A. Wilmstra. He describes how pollen analyses are used to study former climates (and the limitations of such analyses); presents a continent-by-continent review of the recent climatic record inferred from pollen studies (the sequence of vegetation history diagrams is very useful for non-specialists in the

field); and describes autecological and synecological methods for studying past climates from pollen records. He does not, however, discuss how disease and insect devastations of a species can be distinguished from reductions caused by climatic change, particularly in autecological interpretations.

Jean-Claude Duplessy discusses isotope studies in one of the most thorough and well written chapters of the book. He describes the kinds of stable isotope ratios of climatic significance, with particular attention to oxygen isotopes. Considerable attention is given to the proper preparation of foraminifera samples, and how living foraminifera respond to changes in oceanic temperature, pressure and salinity. Duplessy then discusses the long-term isotopic record, using several informative figures, and recounts the historical development of these records and their interpretation.

Climatic change in historical times should have been one of the most readable chapters in the book. Unfortunately the authors wrote such long and poorly organised sentences that I found this chapter the most tedious.

M. I. Budyko discusses the Earth's heat budget, concluding that heat balance calculations can adequately account for the regularities of modern climate. He begins a discussion of irregularities in past climates by introducing elements of catastrophe theory that show how widely different climates can exist for present solar radiation, Earth geography, atmospheric carbon dioxide and aerosol levels. This is the best part of the chapter, as he summarises his own work. Phanerozoic climatic changes due to changes of the above external boundary conditions are also discussed.

Changes in the snow and ice cover have the greatest correlation with the Earth's heat budget. George Kukla discusses these changes in space and time over the past decade, when continuous satellite monitoring began. He starts by describing how the monitoring is done, what areas are covered, and how uncertainties are treated. Then, with the aid of several remarkable figures, he illustrates the effect of seasonal and secular changes in snow and ice, their impact on climate variability and their relationship to solar insolation over the past 130,000 years, and climatic results in both hemispheres. Phase and antiphase relationships are discussed, and correlations with Milankovich orbital variations suggested.

The section on astronomical influences on climate by John Gribbin includes long-term, cyclic and short-term effects. The most interesting long-term effect is treated using the McCrea theory that relates ice epochs, and even ice ages within epochs, to variations in solar luminosity as the Solar System passes through lanes of cold dust between the spiral arms of the Milky Way galaxy.

However, the 250 Myr time of passage doesn't seem to correlate well with ice epochs over the past 2300 million years that Tarling describes in his contribution. Gribbin's discussion of climatic cycles begins by making the point that accumulating random errors can create the impression of cycles where none exist. He then discusses the 100,000, 40,000, and 21,000 yr Minankovich orbital cycles and cites recent work showing climatic variations that correlate with the two shorter cycles. The dominant 100,000 yr ice age cycle is a feature of only the past 600,000 yr or so, suggesting that more than external boundary conditions are involved in triggering ice ages. Again, internal boundary conditions related to ice sheet surges come to mind. But glacial instabilities of this kind are not discussed; neither is the Holocene glacial geological evidence for a 2500 yr climatic cycle reported by Denton and Karlen (1973).

Tim Barnett discusses the ocean's role in the climatic machine. He presents several studies that document the strong linkage of ocean-atmosphere dynamics, but often with significant lag times. This linkage demands that the two be treated as a coupled system. Yet the overwhelming weight of funding goes to atmospheric scientists who treat the ocean as a static boundary condition. The potential contribution of coupled

SCIENTIFIC — BOOKS

H. K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £7.00. (Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

H. K. LEWIS & Co. Ltd.

136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavt,
London, WC1E 6BS."

Circle No. 19 on Reader Enquiry Card.

ocean-atmosphere climatic models is apparent from the discussion by Jill Williams on modelling general atmospheric circulation. Results of four models of the last ice age maximum are compared. Robert Chervin discusses the limitations of general circulation models. He shows how statistical methods are used to separate the signal and noise components of a model's output. An intuitive judgement has to be made in many cases.

The final section of the book discusses how general circulation models can be used to study the interaction of man with today's climate, and what climatic changes this might produce in the future. Will Kellogg examines the effects of increasing land cultivation, fossil fuel

consumption and other industrial contaminants. He concludes that climate will generally become warmer and wetter, especially over the land, and this trend may already be overtaking natural climatic trends. Kellogg suspects that this warming might greatly reduce or even eliminate the arctic pack ice, but sees no reason to think it could affect the stability of present ice sheets until far into the future. Stephen Schneider and Richard Temkin discuss climatic changes and human affairs, particularly food, water, and energy resources.

Terence J. Hughes

Terence J. Hughes is an Associate Professor in the Department of Geological Sciences at the University of Maine, Orono.

Broad survey of diffraction

Diffraction From Materials. By L. H. Schwartz and H. B. Cohen. Pp. 558. (Academic: New York, San Francisco and London, 1977.) \$27.50; £19.95.

DIFFRACTION, especially X-ray diffraction, provides an experimental approach to a wide range of structural problems in such disparate disciplines as chemistry, biology and materials science. The importance of certain aspects of the subject has resulted in the growth of specialisations, and in chemistry, for example, the elucidation of molecular constitution, stereochemistry and geometry by X-ray studies of crystal structure is now a routine procedure with a large body of practitioners and an output of results that requires the Chemical Society Specialist Periodical Report *Molecular Structure by Diffraction Methods* to review more than 2,000 references annually.

There is a risk that scientists immersed in one aspect of diffraction may lose sight of the possibilities that exist for making use of the other aspects of the subject. Accordingly, in addition to the books devoted to the specialised areas of diffraction, there is a place for textbooks that provide a broad survey of the subject. Schwartz and Cohen succeed admirably in this latter aim.

Their book opens with an account of crystal symmetry and then deals with geometrical aspects of crystals and lattices. Stereographic projections are described in some detail in the section on analytical calculations and graphical techniques. Chapter 3 is concerned with kinematical scattering theory, and discussion of diffraction from gratings, rows of atoms and three-dimensional arrays leads to the derivation of the reciprocal lattice and of the

structure factor. Chapters 4 and 5 describe how beams of X-radiation, neutrons and electrons are obtained for diffraction experiments, how these beams interact with matter and how the scattered radiation is detected. The topics covered include absorption and dispersion effects, magnetic scattering of polarised neutrons, and single-crystal and powder diffraction techniques.

The second half of the book develops three major topics. The determination of crystal structures is the theme of chapter 6, in which the authors introduce the principles of Fourier transforms. Electron-density mapping, the Patterson function, intensity statistics, direct phasing procedures and least-squares refinement are clearly explained. Chapter 7 is entitled "What Else Can We Learn from a Diffraction Experiment besides the Average Structure?" and devotes 71 pages to a discussion of thermal diffuse scattering, mosaic size, stacking disorder, local order and clustering, small-angle scattering, and diffraction by liquids and amorphous solids. The final chapter gives an exposition of the dynamical theory of diffraction and concludes with an outline of X-ray topography.

The book is well illustrated and there are useful references to other texts and to original papers. An excellent set of problems follows each chapter, providing the reader with a rigorous test of his mastery of the contents.

Schwartz and Cohen have provided us with one of the best general treatments of the subject currently available and I can recommend their book to all scientists with an interest in diffraction experiments.

George Sim

George Sim is Gardiner Professor of Chemistry at the University of Glasgow, UK.

Valence and molecular orbital theories

Structures and Approximations for Electrons in Molecules. By D. B. Cook. Pp. 294. (Wiley: London, New York and San Francisco, 1978.) £13.50.

THE armoury of books available on valence and molecular orbital theories is so well stocked that some ingenuity is necessary for a new book in this area to be anything other than a re-hash of earlier texts. The mere fact that Dr Cook has succeeded is a credit to his originality, and he convinces the reader that he is producing something quite novel. His earlier book on *ab initio* valence calculations raises expectations which are not disappointed, and no-one could read the book without learning a great deal. The book is particularly important for the many who use semi-empirical molecular orbital methods.

Users of approximate methods are split into two camps: those whose main interest is molecular energetics and computed molecular properties, and those whose chief concern is the interpretation of valence theories. This book follows the second path. Much is made of the theoretical basis of the functional group approach, and the important but ambiguous role of molecular orbitals is clarified in a way not to be found in most books at this level (senior undergraduate and post-graduate students).

A particular strength is the emphasis on computational aspects. This matters, when so many workers use 'off the peg' programmes and only have a vague idea of the inner workings. Encouragement to play with the internal organs of programmes is provided by an Appendix entitled "Invent your own molecular orbital method"; a thought which might not please editors of journals but which should lead to understanding by the chemist.

The book is opinionated and expresses views which will not be universally accepted; but this only adds spice. It is not a polished product and is certainly easier to work from than to read, a conclusion based both on the content and on the production. Justified type is used in the preface and contents but thereafter we have unjustified reproduction of typescript.

Even at a cost above the photocopying limit the book deserves to be bought and used by individuals rather than referred to in libraries.

Graham Richards

Graham Richards is Lecturer in the Physical Chemistry Laboratory at the University of Oxford, UK.